

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 14:14:56 ; Search time 2703 Seconds
(without alignments)
7360.756 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagctctgacgcgcgtctctctctgcatcataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406	100.0	2406	US-10-600-070-1	Sequence 1, Appl1
2	2402.8	99.9	2406	US-10-600-070-9	Sequence 9, Appl1
3	2399.6	99.7	2406	US-10-600-070-128	Sequence 128, App
4	2399.6	99.7	2637	US-10-600-070-130	Sequence 130, App
5	2399.6	99.7	2679	US-10-739-930-227	Sequence 227, App
6	1724	71.7	3667	US-10-600-070-3	Sequence 3, Appl1
7	1720.8	71.5	3667	US-10-600-070-10	Sequence 10, Appl1
8	482	20.0	561	US-10-600-070-132	Sequence 132, App
9	481	20.0	2283	US-10-600-070-126	Sequence 126, App
10	472.8	19.7	1146	US-10-424-599-129007	Sequence 129007, A
11	361.4	15.0	1411	US-10-425-115-81853	Sequence 81853, A
12	328.8	13.7	631	US-10-600-070-184	Sequence 184, App
13	278.8	11.6	660	US-10-600-070-135	Sequence 135, App
14	231.6	9.6	1039	US-10-424-599-35059	Sequence 35059, A
15	224	9.3	537	US-10-600-070-143	Sequence 143, App
16	222.4	9.2	552	US-10-021-323-3636	Sequence 3636, App
17	207	8.6	545	US-10-600-070-175	Sequence 175, App
18	197.2	8.2	491	US-10-600-070-174	Sequence 174, App
19	195.8	8.1	647	US-10-600-070-185	Sequence 185, App
20	193.8	8.1	460	US-10-600-070-187	Sequence 187, App
21	190	7.9	446	US-10-600-070-151	Sequence 151, App
22	188.8	7.8	1703	US-10-437-963-69933	Sequence 69933, A
23	186.6	7.8	652	US-10-600-070-186	Sequence 186, App

24	178.6	7.4	607	US-10-021-323-3562	Sequence 3562, App
25	175	7.3	439	US-09-732-627A-2154	Sequence 2154, App
26	170.4	7.1	608	US-10-600-070-137	Sequence 137, App
27	170.2	7.1	563	US-10-600-070-178	Sequence 178, App
28	169.8	7.1	632	US-10-600-070-146	Sequence 146, App
29	169	7.0	1536	US-10-425-115-57452	Sequence 57452, A
30	167.6	7.0	1032	US-10-767-701-9474	Sequence 9474, App
31	165.4	6.9	527	US-10-600-070-134	Sequence 134, App
32	162	6.7	647	US-10-767-701-4069	Sequence 4069, App
33	158.6	6.6	871	US-10-600-070-153	Sequence 153, App
34	155.4	6.5	307	US-10-600-070-138	Sequence 138, App
35	150.4	6.3	418	US-10-600-070-144	Sequence 144, App
36	146.2	6.1	2130	US-10-437-963-69932	Sequence 69932, A
37	144.6	6.0	653	US-10-600-070-148	Sequence 148, App
38	144.2	6.0	479	US-10-600-070-150	Sequence 150, App
39	130.8	5.4	420	US-10-600-070-176	Sequence 176, App
40	128.6	5.3	480	US-10-600-070-145	Sequence 145, App
41	127.6	5.3	535	US-10-600-070-149	Sequence 149, App
42	124.2	5.2	360	US-10-600-070-179	Sequence 179, App
43	118.8	4.9	336	US-10-600-070-142	Sequence 142, App
44	118.6	4.9	541	US-10-600-070-154	Sequence 154, App
45	117.8	4.9	309	US-10-600-070-141	Sequence 141, App

ALIGNMENTS

RESULT 1
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oeteryoung, Katherine W.
; APPLICANT: Vitha, Stanislaw
; APPLICANT: Kokshtanova, Olga A.
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Query Match	100.0%;	Score 2406;	DB 7;	Length 2406;
Best Local Similarity	100.0%;	Pred No. 0;	Mismatches 0;	Indels 0;
Matches 2406;	Conservative	0;		Gaps 0;
QY	1	ATGGAAGCTCTGAGTACGCTGCGCATTTGCTCTCCCATTTCCATTATGCGATTACCA	60	
DB	1	ATGGAAGCTCTGAGTACGCTGCGCATTTGCTCTCCCATTTCCATTATGCGATTACCA	60	
QY	61	CCGGCAGCAGCAAAAGTCTCGAGTACCAACACCTCTTACACATTTGCTCCGACG	120	
DB	61	CCGGCAGCAGCAAAAGTCTCGAGTACCAACACCTCTTACACATTTGCTCCGACG	120	
QY	121	AAATGGGCGGACCGCTTCTCTCCGCTTCAATTTACCTCGATTCCCTCTCTCC	180	
DB	121	AAATGGGCGGACCGCTTCTCTCCGCTTCAATTTACCTCGATTCCCTCTCTCTCC	180	
QY	181	TTTCGACCGCAGCAGCAGCAGCAGCAGTCTGCTCTCTGACCATTTATGCTGCC	240	
DB	181	TTTCGACCGCAGCAGCAGCAGCAGCAGTCTGCTCTCTGACCATTTATGCTGCC	240	
QY	241	GAAGCGCAGTCCCATCCCATTTGATTTTCAACAGATTTAGAGCTCAACACATTTTC	300	
DB	241	GAAGCGCAGTCCCATCCCATTTGATTTTCAACAGATTTAGAGCTCAACACATTTTC	300	

QY 301 TTAAACGATGGAATCAGAAAGCATTTGCAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 360
| | | | |
Db 301 TTAAACGATGGAATCAGAAAGCATTTGCAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 360
| | | | |
QY 361 TTCAAGGACGACGCTTTTAATCAGCCGAGACAGATTCTTCAGCTGCTGCGAACTCTG 420
| | | | |
Db 361 TTCAAGGACGACGCTTTTAATCAGCCGAGACAGATTCTTCAGCTGCTGCGAACTCTG 420
| | | | |
QY 421 TCTAAATCCGCTCGGTGAAGAGAGTCAATGAAGGCTTCTGATGATGAAGAAAGCTACA 480
| | | | |
Db 421 TCTAAATCCGCTCGGTGAAGAGAGTCAATGAAGGCTTCTGATGATGAAGAAAGCTACA 480
| | | | |
QY 481 GTCAATCACTGATGCTCTTGAGTAAGGTTCTCGGGCTCTCTGTGATTTGCAAGAGGT 540
| | | | |
Db 481 GTCAATCACTGATGCTCTTGAGTAAGGTTCTCGGGCTCTCTGTGATTTGCAAGAGGT 540
| | | | |
QY 541 GGTGAGACTGAGATGATTTCTCGGGTTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAG 600
| | | | |
Db 541 GGTGAGACTGAGATGATTTCTCGGGTTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAG 600
| | | | |
QY 601 TCGTTTAACGAAAGATGCTTTAGTTATGCGGCTTCTGATGCTGAGAGGAT 660
| | | | |
Db 601 TCGTTTAACGAAAGATGCTTTAGTTATGCGGCTTCTGATGCTGAGAGGAT 660
| | | | |
QY 661 GCTATGCAATGATGATCACTGATTTTAATGCTGATGATGATGATGATGATGATGATG 720
| | | | |
Db 661 GCTATGCAATGATGATCACTGATTTTAATGCTGATGATGATGATGATGATGATGATG 720
| | | | |
QY 721 AAGCTTTTACAG 780
| | | | |
Db 721 AAGCTTTTACAG 780
| | | | |
QY 781 GAGACTTTGGAAGAGATCACTCCGCGTATGCTGAGAGCTACTGAGCTTAACGCTTGT 840
| | | | |
Db 781 GAGACTTTGGAAGAGATCACTCCGCGTATGCTGAGAGCTACTGAGCTTAACGCTTGT 840
| | | | |
QY 841 GATGATTAACGCTCGAAGAAAGTAAATGTTAAGCGGTGCGAATATTTTGCTGCT 900
| | | | |
Db 841 GATGATTAACGCTCGAAGAAAGTAAATGTTAAGCGGTGCGAATATTTTGCTGCT 900
| | | | |
QY 901 GTTGAAGAGGTGAGAGATCACTGCTGTTGAGGGGTTGAACCCGTGAAGAGTTATGAT 960
| | | | |
Db 901 GTTGAAGAGGTGAGAGATCACTGCTGTTGAGGGGTTGAACCCGTGAAGAGTTATGAT 960
| | | | |
QY 961 GAGCGCTTTTAAAGATGACAGCTGCTGAGAGAGTTGATCTTTTGAAGTACCCCAAGC 1020
| | | | |
Db 961 GAGCGCTTTTAAAGATGACAGCTGCTGAGAGAGTTGATCTTTTGAAGTACCCCAAGC 1020
| | | | |
QY 1021 AATATTCAGCAGAGTCAATTTGAAGTTTACAGAGTTGCACTTGCTCTTGCTCAAGCT 1080
| | | | |
Db 1021 AATATTCAGCAGAGTCAATTTGAAGTTTACAGAGTTGCACTTGCTCTTGCTCAAGCT 1080
| | | | |
QY 1081 TTTATTTGTAAGAAAGCACAACCTTTTACAGAGTCTGATTAAGCAATTCAGCAACTTCA 1140
| | | | |
Db 1081 TTTATTTGTAAGAAAGCACAACCTTTTACAGAGTCTGATTAAGCAATTCAGCAACTTCA 1140
| | | | |
QY 1141 CAGCGCTAAGGTAATGCTATGAGAGATCTCGCGATGTTGATGATTAACAGGAATTAATG 1200
| | | | |
Db 1141 CAGCGCTAAGGTAATGCTATGAGAGATCTCGCGATGTTGATGATTAACAGGAATTAATG 1200
| | | | |
QY 1201 GAGATGAGCTTGGTCTAAGAAAGGGAATCTGAGCACTGCTTAATAGGCAAGTTGATGAA 1260
| | | | |
Db 1201 GAGATGAGCTTGGTCTAAGAAAGGGAATCTGAGCACTGCTTAATAGGCAAGTTGATGAA 1260
| | | | |
QY 1261 TGCCTATGCTGTTGGCTTGAAGAGAGATTCACATATAGGAATTCAGCTATTTG 1320
| | | | |
Db 1261 TGCCTATGCTGTTGGCTTGAAGAGAGATTCACATATAGGAATTCAGCTATTTG 1320
| | | | |
QY 1321 GAGTTTGTGTTGAGAAATCAATGCTGATGACATGATGATCTCTCTGAGCTATGCAAA 1380
| | | | |
Db 1321 GAGTTTGTGTTGAGAAATCAATGCTGATGACATGATGATCTCTCTGAGCTATGCAAA 1380
| | | | |
QY 1381 TTGTTGAAACCTGTTGAGAGGGGTTGTCTTTCTAGGTTCAAGACACCAAAAGATAA 1440
| | | | |

Db 1381 TTGTTGAAACCTGTTGAGAGGGGTTGTCTTTCTAGGTTAGAGACCAAAAGTAA 1440
| | | | |
QY 1441 AATATTTAACTCGGGGACTACTATGATGATCTGATGTTTGAAGTTTGAAGAGAGT 1500
| | | | |
Db 1441 AATATTTAACTCGGGGACTACTATGATGATCTGATGTTTGAAGTTTGAAGAGAGT 1500
| | | | |
QY 1501 GAGTATGTTCAAGGTTCTCTTAACTGCTGCACTATGAGCAAGATTGAGCCGAG 1560
| | | | |
Db 1501 GAGTATGTTCAAGGTTCTCTTAACTGCTGCACTATGAGCAAGATTGAGCCGAG 1560
| | | | |
QY 1561 CATGTAAGAGCTATGCTATGAGAGAGCTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
| | | | |
Db 1561 CATGTAAGAGCTATGCTATGAGAGAGCTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
| | | | |
QY 1621 AGAACTGGGCTGAACCCAGAGATGCAAGAGACAGTGTATAGTATGATCCTGTTGAT 1680
| | | | |
Db 1621 AGAACTGGGCTGAACCCAGAGATGCAAGAGACAGTGTATAGTATGATCCTGTTGAT 1680
| | | | |
QY 1681 AACATGATGAGCGCTGATGATGAGAGCTGCTGCTTATTTGCAAGACCTGTAAGACCTTCT 1740
| | | | |
Db 1681 AACATGATGAGCGCTGATGATGAGAGCTGCTGCTTATTTGCAAGACCTGTAAGACCTTCT 1740
| | | | |
QY 1741 GAAACTTTGAAACTATATGATTTATGCAATTCAGAGCTGGGCTCTCAGAGAGTACCTGAT 1800
| | | | |
Db 1741 GAAACTTTGAAACTATATGATTTATGCAATTCAGAGCTGGGCTCTCAGAGAGTACCTGAT 1800
| | | | |
QY 1801 GAAACTACTGTTGAAGATGCTCCGCTGATGATGTTAAAGAGAGAGAGAGAGAGAGAG 1860
| | | | |
Db 1801 GAAACTACTGTTGAAGATGCTCCGCTGATGATGTTAAAGAGAGAGAGAGAGAGAGAG 1860
| | | | |
QY 1861 GCTGCTGATGAGCAATGAGATGATTTCACTGTTGAGCCAGAGATTTTCTTAAGAG 1920
| | | | |
Db 1861 GCTGCTGATGAGCAATGAGATGATTTCACTGTTGAGCCAGAGATTTTCTTAAGAG 1920
| | | | |
QY 1921 AGCTCATCTTTTCAACGCAAGATATGATGATGATGATGATGATGATGATGATGATG 1980
| | | | |
Db 1921 AGCTCATCTTTTCAACGCAAGATATGATGATGATGATGATGATGATGATGATGATG 1980
| | | | |
QY 1981 GGGTCACTGAGAGCTGACATTCAGAAAGCACTTCCAGAGATGATGATGATGATGATG 2040
| | | | |
Db 1981 GGGTCACTGAGAGCTGACATTCAGAAAGCACTTCCAGAGATGATGATGATGATGATG 2040
| | | | |
QY 2041 AATATGATATCCAAAGTGGCAAGAAATTAAGTCTGCTGTTTGGGCTGATCAACGCA 2100
| | | | |
Db 2041 AATATGATATCCAAAGTGGCAAGAAATTAAGTCTGCTGTTTGGGCTGATCAACGCA 2100
| | | | |
QY 2101 GAAATGTTACAGAGGTTTGTGATGAGGCGCAATGCTGAAGTTTGAAGTGAAGAGCT 2160
| | | | |
Db 2101 GAAATGTTACAGAGGTTTGTGATGAGGCGCAATGCTGAAGTTTGAAGTGAAGAGCT 2160
| | | | |
QY 2161 GAAATGCTGCAAGAGTGGGTTGATGATTAACCTGTTGAACATATGTTGAAGT 2220
| | | | |
Db 2161 GAAATGCTGCAAGAGTGGGTTGATGATTAACCTGTTGAACATATGTTGAAGT 2220
| | | | |
QY 2221 GTGACAGCTCAGCAGATGGAACCCGTGCTCTGTTGGAAGCACTCTGAGAGAGTCTGCT 2280
| | | | |
Db 2221 GTGACAGCTCAGCAGATGGAACCCGTGCTCTGTTGGAAGCACTCTGAGAGAGTCTGCT 2280
| | | | |
QY 2281 TGTCTATCTGATTTGCTTCAATCCAGAAACATATGCTATGATGATGATGATGATGATG 2340
| | | | |
Db 2281 TGTCTATCTGATTTGCTTCAATCCAGAAACATATGCTATGATGATGATGATGATGATG 2340
| | | | |
QY 2341 AGATAGGAAGTTTCTGCTGCAAGTCAAGGTGGAAGAAATCACTGAAGGCTGTTCTTGA 2400
| | | | |
Db 2341 AGATAGGAAGTTTCTGCTGCAAGTCAAGGTGGAAGAAATCACTGAAGGCTGTTCTTGA 2400
| | | | |
QY 2401 TCATTA 2406
| | | | |
Db 2401 TCATTA 2406
| | | | |

RESULT 2

US-10-600-070-9
; Sequence 9, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-9

Query Match 99.9%; Score 2402.8; DB 7; Length 2406;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAGCTCTGAGTCACTGCGGATGCTCTCTCCCATTCATATGCGCATTAACA 60
DB 1 ATGAAGCTCTGAGTCACTGCGGATGCTCTCTCCCATTCATATGCGCATTAACA 60
QY 61 CCGGCGAGCAAAAGTCTCCGACGAGCAACAACACTCTACAACTATCTGCTCGCCAGC 120
DB 61 CCGGCGAGCAAAAGTCTCCGACGAGCAACAACACTCTACAACTATCTGCTCGCCAGC 120
QY 121 AAATGGGCGGACCGCTCTCTCTCGACTTCAATTTCACTTCGATTCCTCTCTCTCC 180
DB 121 AAATGGGCGGACCGCTCTCTCTCGACTTCAATTTCACTTCGATTCCTCTCTCTCC 180
QY 181 TTGGCCACCGGCAACCAACCGGACCTCTCTCTCTGCGACCACTATATGATCGTCCC 240
DB 181 TTGGCCACCGGCAACCAACCGGACCTCTCTCTCTGCGACCACTATATGATCGTCCC 240
QY 181 TTGGCCACCGGCAACCAACCGGACCTCTCTCTCTGCGACCACTATATGATCGTCCC 240
DB 181 TTGGCCACCGGCAACCAACCGGACCTCTCTCTCTGCGACCACTATATGATCGTCCC 240
QY 241 GAAGCGACGCTCCCATCCCATTTGATTTTCAAGGATTTAGAGGCTCAAAACATTTTC 300
DB 241 GAAGCGACGCTCCCATCCCATTTGATTTTCAAGGATTTAGAGGCTCAAAACATTTTC 300
QY 301 TTAACCGATGAATCAGAAAGCATTCGAAAGTGGGTTTGAACCGCGCAATTCGGT 360
DB 301 TTAACCGATGAATCAGAAAGCATTCGAAAGTGGGTTTGAACCGCGCAATTCGGT 360
QY 361 TTACGAGCAGACGCTTTATCAAGCGGAGACATTTCTTCAAGTCTGCTGCAAACTCTG 420
DB 361 TTACGAGCAGACGCTTTATCAAGCGGAGACATTTCTTCAAGTCTGCTGCAAACTCTG 420
QY 421 TCTAATCTCGGCTAGAGAGAGTCAATGAAGGCTCTTCTGATGATGAAGAGCTACA 480
DB 421 TCTAATCTCGGCTAGAGAGAGTCAATGAAGGCTCTTCTGATGATGAAGAGCTACA 480
QY 481 GTCAATCATGATGTTCTTGGGATTAAGGTTCTGCGGCTCTCTGATGATGAAGAGGT 540
DB 481 GTCAATCATGATGTTCTTGGGATTAAGGTTCTGCGGCTCTCTGATGATGAAGAGGT 540
QY 541 GGTGAAGCTGAGATGATTTCTTGGGATTAAGGTTCTGCGGCTCTCTGATGATGAAGAGGT 600
DB 541 GGTGAAGCTGAGATGATTTCTTGGGATTAAGGTTCTGCGGCTCTCTGATGATGAAGAGGT 600
QY 601 TCGTTTAAAGCAAGATGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TCGTTTAAAGCAAGATGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GCTATGAGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GCTATGAGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

QY 721 AAGCTTTTACAG 780
DB 721 AAGCTTTTACAG 780
QY 781 GAGACTTTGAG 840
DB 781 GAGACTTTGAG 840
QY 841 GATGATTAAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 GATGATTAAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 GTTGAAG 960
DB 901 GTTGAAG 960
QY 961 GAGGCGTTTAAAG 1020
DB 961 GAGGCGTTTAAAG 1020
QY 1021 AATTTTCCAG 1080
DB 1021 AATTTTCCAG 1080
QY 1081 TTTATTTGTAAG 1140
DB 1081 TTTATTTGTAAG 1140
QY 1141 CAGGCTTAAG 1200
DB 1141 CAGGCTTAAG 1200
QY 1201 GAGATTAAG 1260
DB 1201 GAGATTAAG 1260
QY 1261 TGCAGATGAG 1320
DB 1261 TGCAGATGAG 1320
QY 1321 GAGTTTGTGTAAG 1380
DB 1321 GAGTTTGTGTAAG 1380
QY 1381 TTGTTGAAAG 1440
DB 1381 TTGTTGAAAG 1440
QY 1441 AAATTTAAAG 1500
DB 1441 AAATTTAAAG 1500
QY 1501 GAGGTAGTCAAG 1560
DB 1501 GAGGTAGTCAAG 1560
QY 1561 CATGTGAAG 1620
DB 1561 CATGTGAAG 1620
QY 1621 AGAACTCGGCTAAG 1680
DB 1621 AGAACTCGGCTAAG 1680
QY 1681 AACATGTAAG 1740
DB 1681 AACATGTAAG 1740
QY 1741 GAAACTTTGAAAG 1800
DB 1741 GAAACTTTGAAAG 1800
QY 1801 GAAACTTAAG 1860

```

Db      1801  GAAACACTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Qy      1861  GCTGCTGTGTGGCAATTTGACATGATTTCTCTTTCAGCCAGAAATTTTCTTTAAAGC 1920
Db      1861  GCTGCTGTGTGGCAATTTGACATGATTTCTCTTTCAGCCAGAAATTTTCTTTAAAGC 1920
Qy      1921  AGCTCATCTTTTCAAGCAAGATATGATTTCTCTTTCAGCTGATGTCGCTACATA 1980
Db      1921  AGCTCATCTTTTCAAGCAAGATATGATTTCTCTTTCAGCTGATGTCGCTACATA 1980
Qy      1981  GGGTCAGTCAGACCTGACATTCAGAACCACTTCCAGAAATGATGCTAGACTGACAG 2040
Db      1981  GGGTCAGTCAGACCTGACATTCAGAACCACTTCCAGAAATGATGCTAGACTGACAG 2040
Qy      2041  AATATAGTATCCAAAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGTATCACCGATA 2100
Db      2041  AATATAGTATCCAAAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGTATCACCGATA 2100
Qy      2101  GAAATGTTACCAAGGTTTGGATGGGCGAAATGCTAGATTTGAACTGACAGACAGCT 2160
Db      2101  GAAATGTTACCAAGGTTTGGATGGGCGAAATGCTAGATTTGAACTGACAGACAGCT 2160
Qy      2161  GAAACTGCGCAGCTTGGGTTGTTATGATTAATACATGTTGAAACTATCTGTTGACGT 2220
Db      2161  GAAACTGCGCAGCTTGGGTTGTTATGATTAATACATGTTGAAACTATCTGTTGACGT 2220
Qy      2221  GTGACAGTCTCAGCAGATGGAACCCGCTCTGTGTGGAAGCACTGTGAGAGATCTGCT 2280
Db      2221  GTGACAGTCTCAGCAGATGGAACCCGCTCTGTGTGGAAGCACTGTGAGAGATCTGCT 2280
Qy      2281  TGTCTATCTGATTTGGTTTCATCCAGAAACATGCTACTGATGTACAGAACTTACACA 2340
Db      2281  TGTCTATCTGATTTGGTTTCATCCAGAAACATGCTACTGATGTACAGAACTTACACA 2340
Qy      2341  AGATACGAAGTTTCTGTGTCCAAGTCAGGGTGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Db      2341  AGATACGAAGTTTCTGTGTCCAAGTCAGGGTGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Qy      2401  TCATTA 2406
Db      2401  TCATTA 2406

RESULT 3
US-10-600-070-128
; Sequence 128, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Koshmarova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: Use
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patent version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

Query Match      99.7%; Score 2399.6; DB 7; Length 2406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      61  CCGGCGACGACAAAGCTCCGACGTACCAACACCTTACAACTATCTGCTCCGACG 120
Db      61  CCGGCGACGACAAAGCTCCGACGTACCAACACCTTACAACTATCTGCTCCGACG 120
Qy      121  AAATGGGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 180
Db      121  AAATGGGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 180
Qy      181  TTGCGACCGGACCAACGCGCACTCGTCTCTGCGACCATCTATGATCGTCC 240
Db      181  TTGCGACCGGACCAACGCGCACTCGTCTCTGCGACCATCTATGATCGTCC 240
Qy      241  GAAAGCAGTCCCGCATCCCATTTGATTTCAACAGATTTAGAGACTCAACACATTT 300
Db      241  GAAAGCAGTCCCGCATCCCATTTGATTTCAACAGATTTAGAGACTCAACACATTT 300
Qy      301  TTAACGATGGAATCAAGAGCATTCGAAGCTTTCGAACCCGCGCAATTCGCT 360
Db      301  TTAACGATGGAATCAAGAGCATTCGAAGCTTTCGAACCCGCGCAATTCGCT 360
Qy      361  TTCAGGACGAGCGCTTAACTAGCGCGGAGACAGATTTCTCAAGCTGCTGCGAACTCTG 420
Db      361  TTCAGGACGAGCGCTTAACTAGCGCGGAGACAGATTTCTCAAGCTGCTGCGAACTCTG 420
Qy      421  TCTAATCTCTCGCTAGAAAGAGTAAATGAAGTCTTCTTGATGATGAAGAGCTACA 480
Db      421  TCTAATCTCTCGCTAGAAAGAGTAAATGAAGTCTTCTTGATGATGAAGAGCTACA 480
Qy      481  GTCATCATGTATGTTCTTGGGATGAAGTCTGCTCTGCTATGTCAGAAAGGT 540
Db      481  GTCATCATGTATGTTCTTGGGATGAAGTCTGCTCTGCTATGTCAGAAAGGT 540
Qy      541  GGTGAACATGAGATAGTCTGCGGTTGGGAGGCTCTGCTTAAGAGAGTTCCTTAAG 600
Db      541  GGTGAACATGAGATAGTCTGCGGTTGGGAGGCTCTGCTTAAGAGAGTTCCTTAAG 600
Qy      601  TCGTTTAACCAAGATGTGTTTATGATGCGCTTCTGATGTCTGAGAGAT 660
Db      601  TCGTTTAACCAAGATGTGTTTATGATGCGCTTCTGATGTCTGAGAGAT 660
Qy      661  GCTATGCAATGGATCACTGATTTTATTAAGTGTGTTAGAGAGTTCG 720
Db      661  GCTATGCAATGGATCACTGATTTTATTAAGTGTGTTAGAGAGTTCG 720
Qy      721  AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      721  AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy      781  GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTTACCTTACCGCTTGT 840
Db      781  GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTTACCTTACCGCTTGT 840
Qy      841  GATGATTAACGCTCGAAAGAGCTAAATGTTTAAAGCGGTGCGGAATATTTTGTGCT 900
Db      841  GATGATTAACGCTCGAAAGAGCTAAATGTTTAAAGCGGTGCGGAATATTTTGTGCT 900
Qy      901  GTTGAAGAGGTGAGACATCACTCTGTTGGGGGTTTGAACCGGTGGAAGTTATGAAT 960
Db      901  GTTGAAGAGGTGAGACATCACTCTGTTGGGGGTTTGAACCGGTGGAAGTTATGAAT 960
Qy      961  GAGGCGTTTTCAGAAATGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db      961  GAGGCGTTTTCAGAAATGACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy      1021  AATATTCAGACAGATCATTTGAAGTTTACGAAGTTGCACTTCTGTTGGGCTCAAGCT 1080
Db      1021  AATATTCAGACAGATCATTTGAAGTTTACGAAGTTGCACTTCTGTTGGGCTCAAGCT 1080
Qy      1081  TTTATTTGTAAGAGCAACCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db      1081  TTTATTTGTAAGAGCAACCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

```


QY	1141	CAGGCTAAGGTAAATGAGCTATGAGAAATTCCTGGACAGTTGTAATGAATACAGGAATTAATGG	1200
Db	1141	CAGGCTAAGGTAAATGAGCTATGAGAAATTCCTGGACAGTTGTAATGAATACAGGAATTAATGG	1200
QY	1201	GAGATAGACTTCGGCTCTAGAAAGGGGACTCTGTGACTGCTTAATAGGCAAGTTGATGAA	1260
Db	1201	GAGATAGACTTCGGCTCTAGAAAGGGGACTCTGTGACTGCTTAATAGGCAAGTTGATGAA	1260
QY	1261	TGCCGTATGTGGTTGGGCTTAGACAGTGAAGATTCACAATATAGGAATCCAGCTAATGTG	1320
Db	1261	TGCCGTATGTGGTTGGGCTTAGACAGTGAAGATTCACAATATAGGAATCCAGCTAATGTG	1320
QY	1321	GAGTTGTTTTGGAAATTCAAAATGCTATGACATATGATGATCTCCCGACTATGCAAA	1380
Db	1321	GAGTTGTTTTGGAAATTCAAAATGCTATGACATATGATGATCTCCCGACTATGCAAA	1380
QY	1381	TGTGTGAAACCTGTGTGGCAGGGGTTGTCTTCCTAGGTTCAAGACACCAAAAGATATA	1440
Db	1381	TGTGTGAAACCTGTGTGGCAGGGGTTGTCTTCCTAGGTTCAAGACACCAAAAGATATA	1440
QY	1441	AAATTTAAACTCGGGGACTACTATGATGATCCATGATTTTGAAGTTACTTGAAAGAGTG	1500
Db	1441	AAATTTAAACTCGGGGACTACTATGATGATCCATGATTTTGAAGTTACTTGAAAGAGTG	1500
QY	1501	GAGGTAGTCAAGGGTCTCTTACGCTGCTGCAACTATAGGCAAGATTGGACCCGAG	1560
Db	1501	GAGGTAGTCAAGGGTCTCTTACGCTGCTGCAACTATAGGCAAGATTGGACCCGAG	1560
QY	1561	CATGTGAAGAAGCTAGTGTATGCAAGGCACTGCAGAAAGTTTTTCTCCGCTATATACAT	1620
Db	1561	CATGTGAAGAAGCTAGTGTATGCAAGGCACTGCAGAAAGTTTTTCTCCGCTATATACAT	1620
QY	1621	AGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGATCTGTGGT	1680
Db	1621	AGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGATCTGTGGT	1680
QY	1681	AACATATGAGCCCGTGAATGTGTGAGCCTGTCTTTATGCAAGAAAGCTGTAAAGCCCTCT	1740
Db	1681	AACATATGAGCCCGTGAATGTGTGAGCCTGTCTTTATGCAAGAAAGCTGTAAAGCCCTCT	1740
QY	1741	GAAAACTTTGAAACTTAATGATTAATGCAATTCGAGCTGGGGTCTGACAGATAGAGGTTGAT	1800
Db	1741	GAAAACTTTGAAACTTAATGATTAATGCAATTCGAGCTGGGGTCTGACAGATAGAGGTTGAT	1800
QY	1801	GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTAAATCTTA	1860
Db	1801	GAAACTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTAAATCTTA	1860
QY	1861	GCTGCTGCTGTGGCAATTGAGCTGATTTCACTGTTCAAGCAAGATATTTCTTAAAGC	1920
Db	1861	GCTGCTGCTGTGGCAATTGAGCTGATTTCACTGTTCAAGCAAGATATTTCTTAAAGC	1920
QY	1921	AGCTCATCTTTTCAACGCAAGAGATATGTTCTTCTATAGGAATCGATGTGCTACATA	1980
Db	1921	AGCTCATCTTTTCAACGCAAGAGATATGTTCTTCTATAGGAATCGATGTGCTACATA	1980
QY	1981	GGGTCAAGTCAAGCTGACGATTCAGAAACATTTCCCAAGAAATGATGTGATGACCTGACAG	2040
Db	1981	GGGTCAAGTCAAGCTGACGATTCAGAAACATTTCCCAAGAAATGATGTGATGACCTGACAG	2040
QY	2041	AATATATGATTCAGAGTGTGATGAGGAGAAATGCTGAAGATTTTGAATGACAGAGCAGCT	2100
Db	2041	AATATATGATTCAGAGTGTGATGAGGAGAAATGCTGAAGATTTTGAATGACAGAGCAGCT	2100
QY	2101	GAAATGTTAACAGAGTTTTTGAATGGGAGAAATGCTGAAGATTTTGAATGACAGAGCAGCT	2160
Db	2101	GAAATGTTAACAGAGTTTTTGAATGGGAGAAATGCTGAAGATTTTGAATGACAGAGCAGCT	2160
QY	2161	GAAACTGCGCAGCTGGGTGGTTATATATATATACACTGTGAAACTATCTGTGACAGT	2220
Db	2161	GAAACTGCGCAGCTGGGTGGTTATATATATATATACACTGTGAAACTATCTGTGACAGT	2220
QY	2221	GTGACAGTCTCAGCAGATGAAACCCGCTCTGTGTGAAAGCAACTGTGAGAGTCTGCT	2280

```

Db      2221  GCGACAGCTCTCAGCAGATGGAAACCGTGTCTGTGTGGAAAGCACTCTGGAGAGTCTGCT 2280
QY      2281  TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAAGACCTTACACACA 2340
Db      2281  TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAAGACCTTACACACA 2340
QY      2341  AGATACGAAGTTTCTGCGCCAGTCAGGGTGGAAAATCACTGAAAGGCTCTGTTCTTGA 2400
Db      2341  AGATACGAAGTTTCTGCGCCAGTCAGGGTGGAAAATCACTGAAAGGCTCTGTTCTTGA 2400
QY      2401  TCATAA 2406
Db      2401  TCATAA 2406

RESULT 4
US-10-600-070-130
; Sequence 130, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Ostryoung, Katherine W.
; APPLICANT: Vitna, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 130
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-130

Query Match          99.7%; Score 2399.6; DB 7; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1  ATGGAAGCTCTGAGTCACAGTCGGCAATTGGTCTCTCTCCCAATTCGATTAATGCGATTACCA 60
Db      114  ATGGAAGCTCTGAGTCACAGTCGGCAATTGGTCTCTCTCCCAATTCGATTAATGCGATTACCA 173
QY      61  CCGGGAAGAGCAAAACCTCCGAGCTAGGCCAACACACTTACAACTATCTGCTCCGCGAGC 120
Db      174  CCGGGAAGAGCAAAACCTCCGAGCTAGGCCAACACACTTACAACTATCTGCTCCGCGAGC 223
QY      121  AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
Db      234  AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 293
QY      181  TTCGCCACCGCCACCAACACCGGCACCTCTGCTCTCTGCGCACACATCTAATTGATGTCCTC 240
Db      294  TTCGCCACCGCCACCAACACCGGCACCTCTGCTCTCTGCGCACACATCTAATTGATGTCCTC 353
QY      241  GAAGCGCACGTCGCCATCCCATTTGATTTCTACACAGGATTTTGGAGCTCAAAACACATTTTC 300
Db      354  GAAGCGCACGTCGCCATCCCATTTGATTTCTACACAGGATTTTGGAGCTCAAAACACATTTTC 413
QY      301  TTAACCGATGGAATCAGAAAGACATTCGAAAGCTAGGGTTTCCAAAACCGCGCAATTCCGT 360
Db      414  TTAACCGATGGAATCAGAAAGACATTCGAAAGCTAGGGTTTCCAAAACCGCGCAATTCCGT 473
QY      474  TTCAGCGACGACGCTTTATATCAGCCGAGACAGATTTCTCAAGCTGCTTGGCGAAACTCTG 533
Db      421  TTCATATCTCGGTCTCAGAAAGAGATCAATGAAGGCTCTTGTGATGATGAAGAGGTACCA 480
QY      534  TCTATATCTCGGTCTCAGAAAGAGATCAATGAAGGCTCTTGTGATGATGAAGAGGTACCA 593

```

QY 481 GTCACACTGATGTCCTTGGATTAAGGTCCTGAGGCTCTGTGTATGCAAGAGGT 540
 DB 594 GTCATCACTGATGTCCTTGGATTAAGGTCCTGAGGCTCTGTGTATGCAAGAGGT 653
 QY 541 GGTGAGACTGAGATGATCTTCCGGGTGGTGAAGCTCTGCTTAAGAGAGGTCCTTAAG 600
 DB 654 GGTGAGACTGAGATGATCTTCCGGGTGGTGAAGCTCTGCTTAAGAGAGGTCCTTAAG 713
 QY 601 TCGTTTAAGCAAGATGATGTTTAAATGAGCCCTGCTTCCAGATGCTGAGAGGAT 660
 DB 714 TCGTTTAAGCAAGATGATGTTTAAATGAGCCCTGCTTCCAGATGCTGAGAGGAT 773
 QY 661 GCTATGAGCAATGATGATCACTGATTTTAAATGAGCTGCTTAAATGAGCTGCTTAA 720
 DB 774 GCTATGAGCAATGATGATCACTGATTTTAAATGAGCTGCTTAAATGAGCTGCTTAA 833
 QY 721 AACCTTTTACAG 780
 DB 834 AACCTTTTACAG 893
 QY 781 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGGAGCTACTGCTTACCGCTTGT 840
 DB 894 GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGGAGCTACTGCTTACCGCTTGT 953
 QY 841 GATGATTAAGCTGAG 900
 DB 954 GATGATTAAGCTGAG 1013
 QY 901 GTTGAAG 960
 DB 1014 GTTGAAG 1073
 QY 961 GAGGCGTTTTCAGATGAG 1020
 DB 1074 GAGGCGTTTTCAGATGAG 1133
 QY 1021 AATATTCAG 1080
 DB 1134 AATATTCAG 1193
 QY 1081 TTTATGTTAAG 1140
 DB 1194 TTTATGTTAAG 1253
 QY 1141 CAGGCTTAAG 1200
 DB 1254 CAGGCTTAAG 1313
 QY 1201 GAGATGAG 1260
 DB 1314 GAGATGAG 1373
 QY 1261 TGCCTGATGAG 1320
 DB 1374 TGCCTGATGAG 1433
 QY 1321 GAGTTGTTTGAAG 1380
 DB 1434 GAGTTGTTTGAAG 1493
 QY 1381 TTGTTGAG 1440
 DB 1494 TTGTTGAG 1553
 QY 1441 AATATTAAG 1500
 DB 1554 AATATTAAG 1613
 QY 1501 GAGGATGAG 1560
 DB 1614 GAGGATGAG 1673

QY 1561 CATGTGAAGAGTAGTGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 DB 1674 CATGTGAAGAGTAGTGTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1733
 QY 1621 AGAACTCGGCTGAAG 1680
 DB 1734 AGAACTCGGCTGAAG 1793
 QY 1681 AACCAATGAG 1740
 DB 1794 AACCAATGAG 1853
 QY 1741 GAAACCTTTGAAG 1800
 DB 1854 GAAACCTTTGAAG 1913
 QY 1801 GAAACCTTTGAAG 1860
 DB 1914 GAAACCTTTGAAG 1973
 QY 1861 GCTGCTGAG 1920
 DB 1974 GCTGCTGAG 2033
 QY 1921 AGCTCATCTTTTCAAG 1980
 DB 2034 AGCTCATCTTTTCAAG 2093
 QY 1981 GGGTCAGTCAAG 2040
 DB 2094 GGGTCAGTCAAG 2153
 QY 2041 AATATGATTCAG 2100
 DB 2154 AATATGATTCAG 2213
 QY 2101 GAAATGTTACAG 2160
 DB 2214 GAAATGTTACAG 2273
 QY 2161 GAAACCTGCGAG 2220
 DB 2274 GAAACCTGCGAG 2333
 QY 2221 GTGACAGTCAAG 2280
 DB 2334 GTGACAGTCAAG 2393
 QY 2281 TGTCTATCTGATTTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 DB 2394 TGTCTATCTGATTTGCTTCAAG 2453
 QY 2341 AGATACGAAGTTTCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
 DB 2454 AGATACGAAGTTTCTGCTTCAAG 2513
 QY 2401 TCATAA 2406
 DB 2514 TCATAA 2519

RESULT 5
 US-10-739-930-227
 ; Sequence 227, Application US/10739930
 ; Publication No. US20040216190A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
 ; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
 ; FILER REFERENCE: 38-21(53377)B
 ; CURRENT APPLICATION NUMBER: US/10/739, 930
 ; CURRENT FILING DATE: 2003-12-18
 ; NUMBER OF SEQ ID NOS: 11088

SEQ ID NO 227
LENGTH: 2679
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643_1
US-10-739-930-227

Query Match 99.78; Score 2399.6; DB 8; Length 2679;
Best Local Similarity 99.88; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCGAGTCAAGTCCGATGCTCTCCCATTCCTCAATATGCGATTCACA 60
DB 115 ATGGAAGCTCGAGTCAAGTCCGATGCTCTCCCATTCCTCAATATGCGATTCACA 174
QY 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTCAACTATCTGCTCCGACG 120
DB 175 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTCAACTATCTGCTCCGACG 234
QY 121 AAATGGGCGACCGCTCTCTCTCCGACTTCAATTCACCTCCGATCTCTCTCTCTCC 180
DB 235 AAATGGGCGACCGCTCTCTCTCCGACTTCAATTCACCTCCGATCTCTCTCTCTCC 294
QY 181 TTGGCCACCGGCGACCGACCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 240
DB 295 TTGGCCACCGGCGACCGACCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 354
QY 241 GAAAGCGACGCTCCCATCCCATTTGATTTCTACAGGATATTAGAGCTCAAAACATTTTC 300
DB 355 GAAAGCGACGCTCCCATCCCATTTGATTTCTACAGGATATTAGAGCTCAAAACATTTTC 414
QY 301 TTAAACGATGAATCAGAAAGCATTCGAAAGCTAGGCTTTCGAAACCGCGCAATTCGT 360
DB 415 TTAAACGATGAATCAGAAAGCATTCGAAAGCTAGGCTTTCGAAACCGCGCAATTCGT 474
QY 361 TTAGAGGACGACGCTTAAATCAGCGCGGACAGATTTCTCAAGCTGCTTTCGAAACCTG 420
DB 475 TTAGAGGACGACGCTTAAATCAGCGCGGACAGATTTCTCAAGCTGCTTTCGAAACCTG 534
QY 421 TCTAATCTCGCTCTAGAAAGAGTACAAATGAAGCTCTTCTGATGATGAAGAGCTACA 480
DB 535 TCTAATCTCGCTCTAGAAAGAGTACAAATGAAGCTCTTCTGATGATGAAGAGCTACA 594
QY 481 GTCAATCACTGATGCTCTTGGATAGGCTTCGAGGCTCTCTGTGATTTGCAAGAGT 540
DB 595 GTCAATCACTGATGCTCTTGGATAGGCTTCGAGGCTCTCTGTGATTTGCAAGAGT 654
QY 541 GGTGAGCTGAGATGATTTCTCGGCTGATGAGGCTCTGCTTAAAGAGAGGCTTCCAG 600
DB 655 GGTGAGCTGAGATGATTTCTCGGCTGATGAGGCTCTGCTTAAAGAGAGGCTTCCAG 714
QY 601 TCGTTTAAAGCAAGTGTGTTTATGATGAGCGCTTCGTTTCTGATGCTCGAGGGAT 660
DB 715 TCGTTTAAAGCAAGTGTGTTTATGATGAGCGCTTCGTTTCTGATGCTCGAGGGAT 774
QY 661 GCTATGCGATTTGATTCACCTGATTTTATTAAGGCTTATGAGTTTGTGGAAGCTTGG 720
DB 775 GCTATGCGATTTGATTCACCTGATTTTATTAAGGCTTATGAGTTTGTGGAAGCTTGG 834
QY 721 AAATTTTAAAG 780
DB 835 AAATTTTAAAG 894
QY 781 GAGACTTTGGAAGAGATCATCCGCGTATGCTTTGAGGCTACTTGGCTTACCGCTTGT 840
DB 895 GAGACTTTGGAAGAGATCATCCGCGTATGCTTTGAGGCTACTTGGCTTACCGCTTGT 954
QY 841 GATGATTTAGCGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 955 GATGATTTAGCGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
QY 901 GTTGGAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCTGTGAAGAGTTATGAT 960

DB 1015 GTTGGAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCTGTGAAGAGTTATGAT 1074
QY 961 GAGGCGTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
DB 1075 GAGGCGTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCCAAGC 1134
QY 1021 AAATTTCCAGACAGATCATTTGAGTTTAAAGGTTGACATGCTGCTGTTGGGCTCAAGCT 1080
DB 1135 AAATTTCCAGACAGATCATTTGAGTTTAAAGGTTGACATGCTGCTGTTGGGCTCAAGCT 1194
QY 1081 TTTATTTGTAAGAGACACACCTTTTAAAGAGTGTGATTAAGCAATTCAGCAATTCAG 1140
DB 1195 TTTATTTGTAAGAGACACACCTTTTAAAGAGTGTGATTAAGCAATTCAGCAATTCAG 1254
QY 1141 CAGGCTAAGGTAATGCTATGAGAGATCTCTGCAATGTTGATGATACAGGAATTAATGG 1200
DB 1255 CAGGCTAAGGTAATGCTATGAGAGATCTCTGCAATGTTGATGATACAGGAATTAATGG 1314
QY 1201 GAGATGACCTTCCGCTAGAAAGGGGACCTGTGCACTGCTTATAGGCAAGTTGATGAA 1260
DB 1315 GAGATGACCTTCCGCTAGAAAGGGGACCTGTGCACTGCTTATAGGCAAGTTGATGAA 1374
QY 1261 TGCCTATGATGTTGGGCTTGAACAGTGAAGATTCAAAATATAGAAATCCAGCTATTGTG 1320
DB 1375 TGCCTATGATGTTGGGCTTGAACAGTGAAGATTCAAAATATAGAAATCCAGCTATTGTG 1434
QY 1321 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGACATAGCAAA 1380
DB 1435 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGACATAGCAAA 1494
QY 1381 TTGTTGGAACCTGCTTGGAGAGGTTGCTTTCTGATGTTTCAAGGTTCAAGACACAAAGATAA 1440
DB 1495 TTGTTGGAACCTGCTTGGAGAGGTTGCTTTCTGATGTTTCAAGGTTCAAGACACAAAGATAA 1554
QY 1441 AAATTTAACTCCGGGACCTATGATGATCTATGATGATCTATGATGATCTTGAAGAGTGT 1500
DB 1555 AAATTTAACTCCGGGACCTATGATGATCTATGATGATCTATGATGATCTTGAAGAGTGT 1614
QY 1501 GAGGTAGTTCAGGCTTCTCTTAACTGCTGCTGCTGCAATATGCAAGATTCGAGAG 1560
DB 1615 GAGGTAGTTCAGGCTTCTCTTAACTGCTGCTGCTGCAATATGCAAGATTCGAGAG 1674
QY 1561 CATGTGAAGCTAGTGTATGACAGGACCTGCAAAAGTTTCTCCGCTATACAGAT 1620
DB 1675 CATGTGAAGCTAGTGTATGACAGGACCTGCAAAAGTTTCTCCGCTATACAGAT 1734
QY 1621 AGAACTCGGCTGAACCCAAAGATGTCAGAGACAGTGTATGATGATCTGTTGTGT 1680
DB 1735 AGAACTCGGCTGAACCCAAAGATGTCAGAGACAGTGTATGATGATCTGTTGTGT 1794
QY 1681 AACATGTAGCCGTGATGAGCTGCTGCTGCTTATATGCAAGCTGTGAAGCCCTCT 1740
DB 1795 AACATGTAGCCGTGATGAGCTGCTGCTGCTTATATGCAAGCTGTGAAGCCCTCT 1854
QY 1741 GAAAACTTTGAACCTAATGATTTATGCAATTCAGGCTGGGCTCTCAAGAGTACCTGTGAT 1800
DB 1855 GAAAACTTTGAACCTAATGATTTATGCAATTCAGGCTGGGCTCTCAAGAGTACCTGTGAT 1914
QY 1801 GAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
DB 1915 GAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1974
QY 1861 GCTGCTGTGTGCAATTTGACATGATTTCACTGTACAGCCAGAGATATTTTCTTAAAGC 1920
DB 1975 GCTGCTGTGTGCAATTTGACATGATTTCACTGTACAGCCAGAGATATTTTCTTAAAGC 2034
QY 1921 AGCTCATCTTTTCAAGCAAGAGATATGCTTTCTTATGGAATCTGATGTGCTTACATA 1980
DB 2035 AGCTCATCTTTTCAAGCAAGAGATATGCTTTCTTATGGAATCTGATGTGCTTACATA 2094
QY 1981 GGGTCAGTCAAGCTGACATTCAGAGCACTTCCAGAAATGATCTAGACCTGACAG 2040

Db 2095 GGGTCAGTCAGACGTGACGATTCAGAACCACTTCCAGAAATGATGCTAGACTGCAGAG 2154
QY 2041 AATATGATATCCAAAGTGGCAGAAATTAAGTCTTGAGCTTTGGGCTGATCAACCGATA 2100
Db 2155 AATATGATATCCAAAGTGGCAGAAATTAAGTCTTGAGCTTTGGGCTGATCAACCGATA 2214
QY 2101 GAAATGTACAGAGGTTTGGATGAGCGAATGCTGAAATTTGAAGTGAAGAGAGCT 2160
Db 2215 GAAATGTACAGAGGTTTGGATGAGCGAATGCTGAAATTTGAAGTGAAGAGAGCT 2274
QY 2161 GAAACTGCGCAGCTTGAGTTGATGATTAACATGTTGAACATATCTGTTGACAGT 2220
Db 2275 GAAACTGCGCAGCTTGAGTTGATGATTAACATGTTGAACATATCTGTTGACAGT 2234
QY 2221 GTGACAGTCTCAGACAGATGGAACCCGCTCTGCTGAGAACCACTGAGAGAGTCTCT 2280
Db 2335 GTGACAGTCTCAGACAGATGGAACCCGCTCTGCTGAGAACCACTGAGAGAGTCTCT 2294
QY 2281 TGTCTATCTGATTTGGTTCATCCAGAAACATGCTAGTGTGAGAACCTTACACACA 2340
Db 2395 TGTCTATCTGATTTGGTTCATCCAGAAACATGCTAGTGTGAGAACCTTACACACA 2454
QY 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGTTGAAAAATCACTGAGAGCTCTGTTCTTGA 2400
Db 2455 AGATACGAAGTTTCTGCTCAAGTCAAGGTTGAAAAATCACTGAGAGCTCTGTTCTTGA 2514
QY 2401 TCATTA 2406
Db 2515 TCATTA 2520

RESULT 6

US-10-600-070-3
Sequence 3, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Oseeryoung, Katherine W.
APPLICANT: Vitsha, Stanislav
APPLICANT: Kosharova, Olga A.
APPLICANT: Geo, Hong
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 3667
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-600-070-3

Query Match 71.7%; Score 1724; DB 7; Length 3667;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 1 ATGGAAGCTTGAAGTCAAGTGGGATGTTCTCTCCCATTCATTAATGCGATTACCA 60
Db 481 ATGGAAGCTTGAAGTCAAGTGGGATGTTCTCTCCCATTCATTAATGCGATTACCA 540
QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCACAACACCTCTCAACTATCTGCGGAGC 120
Db 541 CCGGCGACGACAAAGCTCCGACGTAAGCACAACACCTCTCAACTATCTGCGGAGC 600
QY 121 AAATGGCCGACGCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
Db 601 AAATGGCCGACGCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
QY 181 TTGGCCACCGCACAACACCGGACGCTCTGCTCTCTGCGACCATTTATGATGCTCC 240
Db 661 TTGGCCACCGCACAACACCGGACGCTCTGCTCTCTGCGACCATTTATGATGCTCC 720

QY 241 GAAAGCCAGCTGCCCATCCCATTTGATCTTACAGATTTAGAGCTCAACACACTTC 300
Db 721 GAAAGCCAGCTGCCCATCCCATTTGATCTTACAGATTTAGAGCTCAACACACTTC 780
QY 301 TTAACCGATGGAATCGAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCCGT 360
Db 781 TTAACCGATGGAATCGAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCCGT 840
QY 361 TTCAGCGACGCTTTTAATCAGCCGAGACAGATTTCTCAAGCTCTTCCGAATCTTG 420
Db 841 TTCAGCGACGCTTTTAATCAGCCGAGACAGATTTCTCAAGCTCTTCCGAATCTTG 900
QY 421 TCTAATCTCGGCTCTAGAAAGATCAATGAAGTCTTCTTGATGATGAAGAGTACA 480
Db 901 TCTAATCTCGGCTCTAGAAAGATCAATGAAGTCTTCTTGATGATGAAGAGTACA 960
QY 481 GTCATCACTGATTTCTCTGGAT----- 504
Db 961 GTCATCACTGATTTCTCTGGATTAAGTAAATTCGATTCGGAATTAAGTTTCTTC 1020
QY 505 -----AGGTTCTCGGAGC 518
Db 1021 GTTTTAATTCATGAATGGATTAAGAAAGAACTTTATCTAGTGAAGTTCTCGGAGC 1080
QY 519 TCTCTGTATTTGCAAGAAAGTGTGAGACTGAGATAGTTCTTCCGCTTGTGAGCTCT 578
Db 1081 TCTCTGTATTTGCAAGAAAGTGTGAGACTGAGATAGTTCTTCCGCTTGTGAGCTCT 1140
QY 579 GCTTAAGAGAGGTTTCCCTTAAGCTTTTAAGCAAGATGTTTATGATGAGCGCTTC 638
Db 1141 GCTTAAGAGAGGTTTCCCTTAAGCTTTTAAGCAAGATGTTTATGATGAGCGCTTC 1200
QY 639 GTTCTGATGTTCCAGAGGATGCTATGAGCATTTGATCCAGCTGATTTTATTAATCTGTA 698
Db 1201 GTTCTGATGTTCCAGAGGATGCTATGAGCATTTGATCCAGCTGATTTTATTAATCTGTA 1260
QY 699 TGAAGTTGTTGAGAAAGCTTTGAAGCTTTTAC----- 730
Db 1261 TGAAGTTGTTGAGAAAGCTTTGAAGCTTTTACAGTAGTTGACTTCTTGTGTAATTC 1320
QY 731 -----A 731
Db 1321 ACGAGGTTGCTTTTAAGAACTTTCTGATTTGATCTTGTATTTAGCTTGTGTA 1380
QY 732 GAGAGAAAGAGCAAGTACCTTGCACCGGATTTACGTGACAATTTAGATGAGACTTGA 791
Db 1381 GAGAGAAAGAGCAAGTACCTTGCACCGGATTTACGTGACAATTTAGATGAGACTTGA 1440
QY 792 AGAGATCACTCCGCTTATGTTTGAAGCTTACCTGCTTACCGCTTGTGATGATTAAGC 851
Db 1441 AGAGATCACTCCGCTTATGTTTGAAGCTTACCTGCTTACCGCTTGTGATGATTAAGC 1500
QY 852 TCGGAAAAGCTTAATGCTTTTAAGCGGTGCGGAAATTTTGTGCTTGTGAGAGAG 911
Db 1501 TCGGAAAAGCTTAATGCTTTTAAGCGGTGCGGAAATTTTGTGCTTGTGAGAGAG 1560
QY 912 TGGAGATCAGCTCTTGTGGGGTTTGAACCGGTGAGAGTTTATGAATGAGCGCTTTT 971
Db 1561 TGGAGATCAGCTCTTGTGGGGTTTGAACCGGTGAGAGTTTATGAATGAGCGCTTTT 1620
QY 972 ACGAATGACAGCTGTGAGC----- 991
Db 1621 ACGAATGACAGCTGTGAGAGATTAACGATTAGATTAATTTTATTTCTTGA 1680
QY 992 -----AGGTTGATCTTTTGTGA 1008
Db 1681 TGAATTAATTTAGGTTTCTCATTTTAATGATGTTGTGATGATGATTTTGTGA 1740
QY 1009 GCTAACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTTGAAGTTGACTTGTCT 1068
Db 1741 GCTAACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTTGAAGTTGACTTGTCT 1800
QY 1069 GTGGCTCAAGCTTTTATGTTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTC 1128

Db	1801	GTGGCTCAAGCTTTTATTGGTAAAGACCAACCTTTTACAGAGATGCTGATTAAGCAATTC	1186
Qy	1129	CAGCAATTCACAGCAGGCTTAAGGTAATGGCTATGAGAATTCCTGCATGTTGATGATACA	1188
Db	1861	CAGCAATTCACAGCAGGCTTAAGGTAATGGCTATGAGAGATTCCTGCATGTTGATGATACA	1920
Qy	1169	CGGATTAATTTGGGAGATGACTTCGGTCTAGAAAGGGGACTCTGTCACCTGTTATAGC	1248
Db	1921	CGGATTAATTTGGGAGATGACTTCGGTCTAGAAAGGGGACTCTGTCACCTGTTATAGGC	1980
Qy	1249	AAAGTTGAATGAAATGCCGTAATGTTGGTGGGCTTAGACAGTGAAGATTCACAAATATAGAAAT	1308
Db	1991	AAAGTTGAATGAAATGCCGTAATGTTGGTGGGCTTAGACAGTGAAGATTCACAAATATAGAAAT	2040
Qy	1309	CCAGCTAATTTGGAGTTGTTTGGAGAAATTCAAATCGTATGACATGATGATCTCCT	1368
Db	2041	CCAGCTAATTTGGAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGAATGATCTCCT	2100
Qy	1369	GGACTATGCAAAATTTGTGGAAACCTGGTGGAGGGGTTGTTCTTCATGGTTACAGAAC	1428
Db	2101	GGACTATGCAAAATTTGTGGAAACCTGGTGGAGGGGTTGTTCTTCATGGTTACAGAAC	2160
Qy	1429	ACCAAGATTAATAAATTTTAACTCCGGGACTACTATGATGATTCCTATGGTTTGGATTAC	1488
Db	2161	ACCAAGATTAATAAATTTTAACTCCGGGACTACTATGATGATTCCTATGGTTTGGATTAC	2220
Qy	1489	TTGGAAGAAGTGGAGTATGTTCAAGGTTCTCCTTAACTGCTGTGCACTATGGCAAG	1548
Db	2221	TTGGAAGAAGTGGAGTATGTTCAAGGTTCTCCTTAACTGCTGTGCACTATGGCAAG	2280
Qy	1549	ATTGGAGCCGAGCATGTAAAGCTATGTCATGACGAGCACTGCAAGAAAGTTTTCCTCC	1608
Db	2281	ATTGGAGCCGAGCATGTAAAGCTATGTCATGACGAGCACTGCAAGAAAGTTTTCCTCC	2340
Qy	1609	CGCTATACAGATAGAAACTCGGCTGAAACCAAGATGTGCAAGAGACAGTGTTAAGTGA	1668
Db	2341	CGCTATACAGATAGAAACTCGGCTGAAACCAAGATGTGCAAGAGACAGTGTTAAGTGA	2400
Qy	1669	GATCTGTTGGTAAATGATAGGCGCTGATGCTGAGCTGGTCTTTATTGCAAGACT	1728
Db	2401	GATCTGTTGGTAAATGATAGGCGCTGATGCTGAGCTGGTCTTTATTGCAAGACT	2460
Qy	1729	GTAAGACCTCTGAAACCTTTGAAACTATGATGTTATGCAATTCGAGCTGGGCTCTCAG	1788
Db	2461	GTAAGACCTCTCTGAAACCTTTGAAACTATGATGTTATGCAATTCGAGCTGGGCTCTCAG	2520
Qy	1789	AGTAGCGTTGATGAAACTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGCAAGT	1848
Db	2521	AGTAGCGTTGATGAAACTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGCAAGT	2580
Qy	1849	GTAAGATTCCTAGCTGCTGCTGTGGCAATTGACCTGATTTCACTGTTCAAGCAAGATAT	1908
Db	2581	GTAAGATTCCTAGCTGCTGCTGTGGCAATTGACCTGATTTCACTGTTCAAGCAAGATAT	2640
Qy	1909	TTTCTTAAGAGCTCATCTTTTCAAGCAAGATATGTTTCTTCATAGGAATCTGAT	1968
Db	2641	TTTCTTAAGAGCTCATCTTTTCAAGCAAGATATGTTTCTTCATAGGAATCTGAT	2700
Qy	1969	GTGCGTACCA-----	1978
Db	2701	GTGCGTACCAATAGATGATTAATAATGATGCAATTTTCATATATCGATTCCTCAATAA	2760
Qy	1979	-----	1978
Db	2761	TGCTTGTTTGTGACCTAAGAACATATGTTCCACTTAATACATGTCCCAAAAGTTGTAAC	2820
Qy	1979	-----	1978
Db	2821	AAAGATTAACAAGTCTGAGTAAATTTTCACTAATTAATGCTGCTGAATTTTTCATCAAA	2880
Qy	1979	-----	1978

[illegible]

QY 61 CCGGAGAGCAAAAGCTCCGAGTAGCCACAACGCTTACAATCTGCTCCGCGAGC 120
DB 541 CCGGAGAGCAAAAGCTCCGAGTAGCCACAACGCTTACAATCTGCTCCGCGAGC 600
QY 121 AATGAGGCGAGCGCTTCTCTCGAATTCAATTTGACCTCGGATTCCTCTCTCC 180
DB 601 AATGAGGCGAGCGCTTCTCTCGAATTCAATTTGACCTCGGATTCCTCTCTCC 660
QY 181 TTGCGCACCGCCACACACCGGCACTCTGCTCTCTCGACACATCTATTTGATCGTCC 240
DB 661 TTGCGCACCGCCACACACCGGCACTCTGCTCTCTCGACACATCTATTTGATCGTCC 720
QY 241 GAAAGCGAGTCCCGCATCCCATTTGATTTACAGGATTTAGAGCTCAACACATTC 300
DB 721 GAAAGCGAGTCCCGCATCCCATTTGATTTACAGGATTTAGAGCTCAACACATTC 780
QY 301 TTAAACGATGAAATCAGAAAGAGATTCAGAGCTAGGATTCGAAACCGCGCAATTCG 360
DB 781 TTAAACGATGAAATCAGAAAGAGATTCAGAGCTAGGATTCGAAACCGCGCAATTCG 840
QY 361 TTACGAGAGAGCGCTTTATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
DB 841 TTACGAGAGAGCGCTTTATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 900
QY 421 TCTAATCCTCGGCTGAAGAGATCAATGAAGGCTCTTGTGATGATGAAGAGCTACA 480
DB 901 TCTAATCCTCGGCTGAAGAGATCAATGAAGGCTCTTGTGATGATGAAGAGCTACA 960
QY 481 GTCATCACTGATGCTCTTGAGAT-----AGGTTCTCTGAGGCT 504
DB 961 GTCATCACTGATGCTCTTGAGATAGGTAATTCGATTCGGAATTAATTAAGTTCTTC 1020
QY 505 -----AGGTTCTCTGAGGCT 518
DB 1021 GTTTAATTTCAATGAATGATTAAGAGAGAACTTTTATCTAGTAGAGGTTCTCTGAGGCT 1080
QY 519 TCTCTGTGATTTCAAGAGAGGAGTGTGAGCTGAGATAGTCTTCCGAGTGTGAGGCTCT 578
DB 1081 TCTCTGTGATTTCAAGAGAGGAGTGTGAGCTGAGATAGTCTTCCGAGTGTGAGGCTCT 1140
QY 579 GCTTAAGAGAGGTTGCTTAAGTCGTTTAAAGCAAGATGTGTTTAACTTTAATGCGGCTTC 638
DB 1141 GCTTAAGAGAGGTTGCTTAAGTCGTTTAAAGCAAGATGTGTTTAACTTTAATGCGGCTTC 1200
QY 639 GTTTCCTGATGCTCGAGGAGTGTATGCAATTCGATTCGATCTTATTTATCTGCTTA 698
DB 1201 GTTTCCTGATGCTCGAGGAGTGTATGCAATTCGATTCGATCTTATTTATCTGCTTA 1260
QY 699 TGAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 730
DB 1261 TGAGTTTGTGAGAGCTTTGAAGCTTTTACAGGTAAGTGTGACTTGTGTAATTTG 1320
QY 731 -----A 731
DB 1321 ACGAGCTTGCGCTTATTAAGAACTTCTGATTTGATTACTTTGTTATTTGAGCTTGTGTA 1380
QY 732 GAGAGAGAGAGCAATGACCTTGTGACCGGATTTACGTGCACAAATTTGATGAGATTTGGA 791
DB 1381 GAGAGAGAGAGCAATGACCTTGTGACCGGATTTACGTGCACAAATTTGATGAGATTTGGA 1440
QY 792 AGAGATACCTCCGCTTATGCTTGAAGCTACTTGGGCTTACCGCTGCTGATGATTAACGC 851
DB 1441 AGAGATACCTCCGCTTATGCTTGAAGCTACTTGGGCTTACCGCTGCTGATGATTAACGC 1500
QY 852 TGCGAAAAGACTAATGATGTTTAAAGCGGTGCGGAAATTTTGTGCTGTGTGAGAGAG 911
DB 1501 TGCGAAAAGACTAATGATGTTTAAAGCGGTGCGGAAATTTTGTGCTGTGTGAGAGAG 1560
QY 912 TGAGAGATCAGCTCTTGTGTGGGGTTTGACCGGTGAGAGATTTATGAATGAGGGGTTTT 971
DB 1561 TGAGAGATCAGCTCTTGTGTGGGGTTTGACCGGTGAGAGATTTATGAATGAGGGGTTTT 1620
QY 972 ACGAATGACAGCTGCTGAG----- 991

DB 1621 ATGAATGACAGCTGCTGAGCAGATATACAGTTAGATACCTTTTAAATTTCTTTAGCA 1680
QY 992 -----AGGTTGATCTTTTGTGA 1008
DB 1681 TGAATATATTAGGTTTCTCATTTTAAATGATGTTGTGTGTGATGATCTTTTGTGA 1740
QY 1009 GCTACCCCAAGCAATATTCACAGAGTCAATTTGAAGTTTACGAAATGTCACCTGCTCTT 1068
DB 1741 GCTACCCCAAGCAATATTCACAGAGTCAATTTGAAGTTTACGAAATGTCACCTGCTCTT 1800
QY 1069 GTGGCTCAAGCTTTTATTTGTTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTC 1128
DB 1801 GTGGCTCAAGCTTTTATTTGTTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTC 1860
QY 1129 CAGCACTTACAGAGCTTAAGGTAATGCTATGAGAGATTCGCGAGTTGTATGATACA 1188
DB 1861 CAGCACTTACAGAGCTTAAGGTAATGCTATGAGAGATTCGCGAGTTGTATGATACA 1920
QY 1189 CCGAATTAATTTGGAGATAGACTTCGCTTGAAGAGGGACTGTGCACTGCTTATAGGC 1248
DB 1921 CCGAATTAATTTGGAGATAGACTTCGCTTGAAGAGGGACTGTGCACTGCTTATAGGC 1980
QY 1249 AAAATTGATGATGCGGTATGTGCTTGGCTTGAAGAGGATTTCAATTAATTAAGAT 1308
DB 1981 AAAATTGATGATGCGGTATGTGCTTGGCTTGAAGAGGATTTCAATTAATTAAGAT 2040
QY 1309 CCAAGCTATGTTGAGATTTGTTTGAAGATTCGAATCGATGACAAATGATGATCTCCCT 1368
DB 2041 CCAAGCTATGTTGAGATTTGTTTGAAGATTCGAATCGATGACAAATGATGATCTCCCT 2100
QY 1369 GGAATATGCAAAATTTGTTGGAACCTGTGTCAGAGGATTTCTTCTAGTTCAAGAGAC 1428
DB 2101 GGAATATGCAAAATTTGTTGGAACCTGTGTCAGAGGATTTCTTCTAGTTCAAGAGAC 2160
QY 1429 ACCAAAGATTAATAATTTTAACTCGGGGACTATGATGATCTATGTTTGAATTC 1488
DB 2161 ACCAAAGATTAATAATTTTAACTCGGGGACTATGATGATCTATGTTTGAATTC 2220
QY 1489 TTGGAAGAGTGAAGGATGCTGAGGTTTCTCTTATGCTGCTGCAACTATGAGCAAG 1548
DB 2221 TTGGAAGAGTGAAGGATGCTGAGGTTTCTCTTATGCTGCTGCAACTATGAGCAAG 2280
QY 1549 ATTGAGCCGAGCATGTGAAGCTAGTGTATGCAAGCACTGAGAAAGTTTCTCTCC 1608
DB 2281 ATTGAGCCGAGCATGTGAAGCTAGTGTATGCAAGCACTGAGAAAGTTTCTCTCC 2340
QY 1609 CCGTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGTA 1668
DB 2341 CCGTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGTA 2400
QY 1669 GATCCTGTTGTTAACAATGTAAGCCGTAAGGATGTAAGCTGCTTATTTGCAAGAGCT 1728
DB 2401 GATCCTGTTGTTAACAATGTAAGCCGTAAGGATGTAAGCTGCTTATTTGCAAGAGCT 2460
QY 1729 GTAAGACCTCTGTAATACTTTGAACATAATGATTAATGCAATTCGAGCTGCGGCTCTAG 1788
DB 2461 GTAAGACCTCTGTAATACTTTGAACATAATGATTAATGCAATTCGAGCTGCGGCTCTAG 2520
QY 1789 AGTAGCGTTGATGAACCTAGTTGAATATGCTGCTGCTATATGTTTAAAGAGGCAAGT 1848
DB 2521 AGTAGCGTTGATGAACCTAGTTGAATATGCTGCTGCTATATGTTTAAAGAGGCAAGT 2580
QY 1849 GTGAAGATCTAGCTGTGTTGAGGCAATTTGGACTGATTTCACTGTTCAAGCAAGATAT 1908
DB 2581 GTGAAGATCTAGCTGTGTTGAGGCAATTTGGACTGATTTCACTGTTCAAGCAAGATAT 2640
QY 1909 TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGGATATGTTTCTTATGGAATCTGAT 1968
DB 2641 TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGGATATGTTTCTTATGGAATCTGAT 2700
QY 1969 GTGCTTACCA----- 1978

```

Db      2701 GTCGCTACCATAGTATGATTAATGATGCAATTTTCATATATCTGCATGCTCAATA 2760
      1979 ----- 1978
Db      2761 TCGCTGTTTGTGAGCTAAGAACATAGTCCCACTTAATACATGTCCCAAAAGTTGACC 2820
      1979 ----- 1978
Db      2821 AAGATTAAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGTAATTTTTCATCAA 2880
      1979 ----- 1978
Db      2881 CTGTAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTAGATAAGTAGATTAAG 2940
      1979 ----- 1979
      1979 -----T 1979
Db      2941 AGATGCTTAGTGTGCTTTGTCACATTTTCTTCTGTGATTTTCTTTTCGATTT 3000
      1980 AGGCTCAGTACAGACTGACGATTTGACAGACACTTCCAGAAATGAGTGTAGACTGACAGA 2039
      3001 AGGCTCAGTACAGACTGACGATTTGACAGACACTTCCAGAAATGAGTGTAGACTGACAGA 3060
      2040 GAATATAGTATCCAAAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 2099
      3061 GAATATAGTATCCAAAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 3120
      2100 AGAAATTTTACCAG----- 2113
      3121 AGAAATTTTACCAGAGGTGAGGGAATTAATCTACAAATTCATCAATTCATCAATTCATCA 3180
      2114 -----AGTTTTTGGATGGG 2128
      3181 TGCAATGATTAATAGTCTGCTGCTTGTGATTTCTGTATTAATAGATTTTGGATGGG 3240
      2129 GAATGCTGAAGATTGGAAGTGAAGACAGACAGCTGAAGCTGGCAGCTTGGTGGTTATG 2188
      3241 GAATGCTGAAGATTGGAAGTGAAGACAGACAGCTGAAGCTGGCAGCTTGGTGGTTATG 3300
      2189 AATTATACCTGTTGAACTATCTGTGTCAGAGTGAAGTCTGACGAGTGAAGAAACCCGCTG 2248
      3301 AATTATACCTGTTGAACTATCTGTGTCAGAGTGAAGTCTGACGAGTGAAGAAACCCGCTG 3360
      2249 CTCTGCTGAAGCAACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTCATCCAGAAA 2308
      3361 CTCTGCTGAAGCAACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTCATCCAGAAA 3420
      2309 ACAATGCTACTGATGTCAGAACCTACACAACAAGATACGAAGTTTCTGTGTCAGAGTCA 2368
      3421 ACAATGCTACTGATGTCAGAACCTACACAACAAGATACGAAGTTTCTGTGTCAGAGTCA 3480
      2369 GGTGGAATCACTGAAGGCTCTGTTTTCGATCATTA 2406
      3481 GGTGGAATCACTGAAGGCTCTGTTTTCGATCATTA 3518

```

RESULT 8
US-10-600-070-132/c

```

; Sequence 132, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseeryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placelid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; LENGTH: 561

```

```

; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (127)..(127)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (520)..(520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-132

```

Query Match 20.0%; Score 482; DB 7; Length 561;
Best Local Similarity 99.4%; Pred. No. 1e-134;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

      1922 GCTCATCTTTTCAACGGAAGATTAAGTCTCTCTATGGAATCTGATGCTGCTACCAATAG 1981
      561 GCTCATCTTTTCAACGGAAGATTAAGTCTCTCTATGGAATCTGATGCTGCTACCAATAG 502
      1982 GGTACGTCAGAGCTGACGATTTCAAGACACTTCCAGAAATGAGTGTAGACTGACAGAGA 2041
      501 GGTACGTCAGAGCTGACGATTTCAAGACACTTCCAGAAATGAGTGTAGACTGACAGAGA 442
      2042 AATATGATTCAGAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 2101
      441 AATATGATTCAGAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 382
      2102 AATATGATTCAGAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 2161
      381 AATATGATTCAGAGTGGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 322
      2162 AACTGCGCAGCTTGGGCTGTATGATTAATCACTGTTGAATCTGTTGACAGCTG 2221
      321 AACTGCGCAGCTTGGGCTGTATGATTAATCACTGTTGAATCTGTTGACAGCTG 262
      2222 TGAAGTCTGAGCAGATGGAACCCGCTGCTGTGTCAGAGCACTGTGAGAGTCTGCTT 2281
      261 TGAAGTCTGAGCAGATGGAACCCGCTGCTGTGTCAGAGCACTGTGAGAGTCTGCTT 202
      2282 GTCTATCTGATTTGTTGTTCACTCCAGAAACAAATGCTACTGATGTCAGAACTTACACAA 2341
      201 GTCTATCTGATTTGTTGTTCACTCCAGAAACAAATGCTACTGATGTCAGAACTTACACAA 142
      2342 GATACGAAGTTTCTGGTCCAGATCAGGAGGAAATCACTGAAGGCTCTGTTCTTGCAT 2401
      141 GATACGAAGTTTCTGGTCCAGATCAGGAGGAAATCACTGAAGGCTCTGTTCTTGCAT 82
      2402 CATTA 2406
      81 CATTA 77

```

RESULT 9
US-10-600-070-126

```

; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseeryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kosharova, Olga A.
; TITLE OF INVENTION: Placelid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2

```


SEQ ID NO 126
LENGTH: 2283
TYPE: DNA
ORGANISM: *Oryza sativa*
US-10-600-070-126

Query Match 20.0%; Score 481; DB 7; Length 2283;
Best Local Similarity 54.3%; Pred. No. 5.2e-134;
Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
OY 108 CTGCTCGCCAGCAATAGGCGCCGACCTCTTCTCTCCGACTTCAATTTCACTCCCATTC 167
DB 132 CTGGCGCGAAGCGCTCTTCCCGACTTCCACCTCTCCCAACCGCGCCGCTCCGACCC 191
OY 168 CTCTCTCTCTCTCTGCGCACCGGCACACCGACCTCTGCTCTGCGCACCATC 227
DB 192 GCGGTCCCGGCGCGCGCGCGCGCGCGCGCGCTCCGCTCCCGCTTCCGCTCTT 251
OY 228 TATTGATCGTCCGAAACGCCACGTCCTCCCATCCCATTTGATTTCTACCGATTATAGAGC 287
DB 252 CCGCGACGCGCGGAAAGCTCCCTCCGCTCCAAAGTGAATTTCTAAGAGTTCTAGGGGC 311
OY 288 TCAACACATTTCTTAAACCGATGGAATCAGAAAGCATTTGAAAGCTAGGGTTTCAACC 347
DB 312 AGAGCCACATTTCTTGGCGATGCGATCAGAGAGGCGTTGAGAGCCAGATAGCCAAAGCC 371
OY 348 GCGGCAATTGCGTTTCAAGGACGACGCTTAACTACGCGGAGACAGATTTCTGAAGTGC 407
DB 372 ACCGCAATATGCTACAGACGAGATGCTCTTGTGTGCTGACAAATATGTCAGATTGC 431
OY 408 TTGCGAACTCTGCTATCTCTCGTCTAGAAAGAGATCAATGAAAGTCTTCTTATGA 467
DB 432 CCAATGACCTCTATGAAACAGAACTCCGCACTCAATGATGATGCTGCTTCTGAGAA 491
OY 468 TGAAGAAGCTACAGTCACTGATGTTCTTGGGATAGAGTTCCTGGGCTCTCTGTGT 527
DB 492 CCGTGAAGAAGCTCTCACTGATGATGCTTGGGAAAGAGAGCTGGG----- 540
OY 528 ATTGCAAGAAGGTGTGAGCTGAGATGTTCTTCGGGTTGTGAGGCTCTGCTTAAGA 587
DB 541 -----GAGGCACTTGTCTGTGTGTTGTAACCTGAGAAACAGTTGCTTCTGGA 584
OY 588 GAGGTTGCTTAAGTCGTTTAAAGCAAGATGTGTTTAAAGTGGCGCTTCCGTTTCTGA 647
DB 585 TCGGCAACCCAGCGCTTCAAGCAGAGACGTGTGCTAGCATGCTCTGCTTAATGTGA 644
OY 648 TGTCTCGAGGATGCTATGSCATTTGATTCACCTGATTTTATTACTGTTATGAGTTGT 707
DB 645 TCTATCAAGGAGTGTCTATGSCAGAACCTCCAGATGTAATGGCTGTGCGAGTGT 704
OY 708 TGAAGAAGCTTTGAAGCTTTTACAGAGAAAGAGCAAGTACCTTGCACCTGAATTCGT 767
DB 705 CGAGAGGCTCTCAAGCTCTTGCAGAGAAAGTGAAGCAAGTCTGCACCTGAATTCGT 764
OY 768 TGCAGAAATTTGAGACTTTGAAAGATCACTCCGCTTATGTTCTTGAAGCTACTTGG 827
DB 765 TTCAACAGATTGAGAACTCTCAAGAGATTTACCTCGCTGTATTTGAAGCTCTCTC 824
OY 828 CTTACCGCTTGTGATGATTAACGCTGCGAAAGACTAAATGTTTAAACGCTGTGCGAA 887
DB 825 CTTTCTTATTAAGACAGAGATCATTAAGAAAGCGCAAGAGGCTTCAAGGTGAGAA 884
OY 888 TATTTTGTGTCTGTTGAGAGAGTGAAGATCAAGCTCTTGTGTGGGGTTTGAACCGTGA 947
DB 885 CATTTGTGTGAGGCTTGTGAGAGAGGATTTGTCTACCGTTGAGAGAGATTTTCTCTGA 944
OY 948 GAGTTTAAATGAGGCGTTTAAAGATGACAGCTGCTGAGGAGGTTGATCTTTTGT 1007
DB 945 AGCTTCATTAAGAGGCTTTTGTGAGAGATCACTCAATTGAATGATGATTTTCTTTC 1004
OY 1008 AGCTACCCCAAGCAATATTCAGAGAGATCAATTGAAGTTTACAGAGTTGCACTTGTCT 1067
DB 1005 AAAAAACCGAATAGCATTTCTCTGATGTTGAAATTTTACAAATGATAGCACTTGCACA 1064
```

```
OY 1068 TGTGCTCAAGCTTTTATTTGTAAGAGCCACACTTTTACAGAGATGCTGATAGCAATT 1127
DB 1065 TGTGCTCAAGCAATTAATGATTAAGAGCCACAAATTCATCATGATGCGGATGATCTTTT 1124
OY 1128 CCAAGCACTTCAAGCAGGCTTAAGGTAATGCTATGAGATTTCTCGCATTTGTATGATAC 1187
DB 1125 TGAACAACTCCAGAAATTTCA-----CATAGTTCTCATTA 1160
OY 1188 ACGAATAATTTGGAGATAGACTTGTGCTTAAAGAGGAGACTCTGTGACCTCTTATAG 1247
DB 1161 TGTATATGATTAATGATGATGACTTGTGATTTGAAAGGAGATTTCTGTCTATTTGATCG 1220
OY 1248 CAAAGTTATGATGAGTGGATGATGTTGGGCTTAAACAGTGAAGATTTCAATATAGAA 1307
DB 1221 AGATTTTACAAAGTGAAGATGCTGATGATTTGATTAATGATTTTCACTTACAGAGA 1280
OY 1308 TCCAGCTATTTGAGAGTTGTTTGGAGAAATTCAAATC--GTGATGACAAATGATGATCT 1364
DB 1281 CCCCAAAATTTAGAGTTTATTTGTGACCAACTCTAGCATCAAGTGAAGAAATGATCTCT 1340
OY 1365 CCTGGAATATGCAATTTGTGAAACCTGTGTGCGAGGGGTTGTCTTCTAGTTTCAG 1424
DB 1341 TCCAGGCTGTGCAAGCTTTTGAAGCTTGTCTTATCTTTGAGTTTCTTAGAGACAG 1400
OY 1425 AGACACCAAGATTAATAAATTTAACTCGGAGACTATATGATGATCTATGTTTGAAG 1484
DB 1401 AGATATCGGGGATATGACATTTGAGACTTGTCTTATCTTTGAGTTTCTTAGAGACAG 1460
OY 1485 TTACTTGAAGAAGTGAAGTGAAGTCAAGGCTTCTCTTTAGCTGTCTGCTCAACTATGCG 1544
DB 1461 CTACCTTGAAGAAGATGAGAGTGTGTGCTCTCTCAATTTGTGCTGTCTGCTCTATATTC 1520
OY 1545 AAGGATTTGAGCGGACATGTGAAGCTATGTGTATGACAGGACCTGCAGAAAGTTTTC 1604
DB 1521 AAAACTTGTGTCTCAAGCTACAGCTGCACTTGTGTA----- 1555
OY 1605 TTCCGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGATGTTTGA 1664
DB 1556 -----CTGTGAATCAAAATGCTATTCAGAGCG----- 1581
OY 1665 TGTAGATCTGTTGTAACAATGTAAGCCGTGATGCTGATGCTGTTCTTATTTGCAGA 1724
DB 1582 -----TTCAACAA 1589
OY 1725 AGCTTAAGACCTCTGAAACCTTTGAACCTATATGATTTATGCAATTCAGCTGGGCTTC 1784
DB 1590 GGTTTTTCATTTGATGAACAGTTAGACAGTCAAGCTAGCAAGAAATATCTAAGATGGCC 1649
OY 1785 AGAGATGAGCTTGAATGAACCTACTGTGAAATGTCGTTGCTGATATGTTAAAGAGCG 1844
DB 1650 TGGGGATATCTTGAATAATTTTGAACAGAGAAATGACCTGCTCATGATTCGAAATG 1709
OY 1845 AAGTGAAGATCTAGCTGTGTGTGCAATTTGATGATTTTCACTGTTTCAAGCCAGAA 1904
DB 1710 CGCTTGAAGATTAATCTGTGTGCGCACTGTTTGAAGCTTGGCAGTAATTTGGGCGCAA 1769
OY 1905 GTATTTTCTTAAAGAGCTCATCTTTTCAACCAAGATATGTTCTTCTATGGAATC 1964
DB 1770 ATATTTGCTCTGTGAAGGCGCTTCTGTCTATTTAAGAGTGAAGATGATCTGTGCAAGT 1829
OY 1965 TGAT-----GTGCTAACATAGGCTCAGTCAAGCTGACATTCAGAAAGCACTTCC 2015
DB 1830 TGTATTAATGATGATCTTATCTGATATCTGTGACATAGATGAAAGTCCAGTATATTC 1889
OY 2016 CAGATGATGCTTAGATGCTGACAGAGAAATATGATATCAAGTGCAGAGATTAAGTCTCT 2075
DB 1890 TAAATGATGAGAGCTGAGAGATATGTTTCCAGTGGCAGATATCAAACTTAA 1949
OY 2076 GGTCTTTGGGCTGATCACCGCATTAAGATTTTACAGAGGTTTGAATGGGGAATGCT 2135
DB 1950 GGCCTTGGGACAGAACATTTGCTGATCATCATTTGCAAGAGTTCCTTGAATGGCAACATGCT 2009
```

```

QY 2136 GAAGATTGGACGTACAGACGAGCTGAAGTCTGCGACCTTGGGTGTTATGATTATAC 2195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2010 AAGGTGTGACGTACGACGAGGAGATTGAGCGTTCATGGGTGTTGGAGTATAC 2069
QY 2196 ACTGTGAAACATCTTGTGACAGTGTGACAGTCTGACAGATGAAACCGGTGCTGTG 2255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2070 ACTATCGAGTGTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 2129
QY 2256 GAGAGCACTCTGAGAGAGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2130 GAGAGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2189
QY 2316 TACTGATGTCAGAACCTACACACACACACACACACACACACACACACACACACAC 2372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2190 TTCTATGATGACAAATATACACTACCGGTGATGATGATGATGATGATGATGATGATG 2249
QY 2373 GAAATATCTGAAGGCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2250 GAAGATTAACGAGAGACGATCTCTCAAGTCGTA 2282

RESULT 10
US-10-424-599-129007
; Sequence 129007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129007
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87499C.1
US-10-424-599-129007

Query Match 19.7%; Score 472.8; DB 7; Length 1146;
Best Local Similarity 70.5%; Pred. No. 1e-131;
Matches 649; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 107 TCTGCTCCGCGACGAATGAGCGGACCGTCTCTCTGCACTTCAATTTGACCTCCGATT 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 TATCCGCGACGAAATGAGCGGACCGTCTCTCTGCACTTCAATTTGACCTCCGAGC 295
QY 167 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 CCGCGCGCTCACTT-----CGACCTCACTCTCAAGCCCTCTCTCTCTCTCTCTCT 346
QY 227 CTATTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 GCCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 CTCAACACATTTCTTAACGATGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 CCGAACCGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 347 CGCGCAATTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 CTCTCTAGTACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 407 CTTCGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 527 CCTCGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 467 ATGAAGAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 587 ACGAAGAGCGCGCATCTCTCACTCAATCCCTTTCGACAAAGTCTCTGAGCGTGTGTG 646
QY 527 TATTGCAAGAGGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATG 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 647 TGTTCGAGAGAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 AGAGGTGCTTAAGTCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATG 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 707 AGAGGTGCTTAAAGTCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATG 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 647 ATGCTCGAGAGATGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATG 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 767 ACGTCTCAAGGATGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATG 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 TTGAGAGAGCTTTGAGAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 827 TCGAGAGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 767 GTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 887 AAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 827 GCTTACCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 947 CTTGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 887 ATATTTTGTGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1007 ACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1066
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 947 AGAAGTTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1067 AAGCTTCTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1007 TAGCTACCCCAACCAATATT 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1127 TAGCCACACCAAGTACTATT 1146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-425-115-81853
; Sequence 81853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81853
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174655C.1
US-10-425-115-81853

Query Match 15.0%; Score 361.4; DB 8; Length 1411;
Best Local Similarity 61.3%; Pred. No. 7.4e-98;
Matches 600; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

QY 96 CTCTAACAATCTGCTCCGCGACGAATGAGCGGACCGTCTCTCTCTCTCTCTCTCTCT 155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 CCCCTCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 156 CACCTCGATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 271 COTCCCC --- CGCGCGCGAGCCCGGCGAGCGCGGCGCTCCTCTTCCCTCTGCTCCCGGT 327

Qy 216 TCTGCCACCATCTATTGATGCTGCCGAGCGCCAGCTGCCATGCCATTGATTTCTACCA 275

Db 328 CGTCCCGATCTTCCCGGAGCCCGCCAGCCGCGCTTGCCCCCTCCCGGTGACATTTCTACAA 387

Qy 276 GGTATTAGAGGCTCAACAACATTTCTTAAACGATGGAATGAAAGAAAGACATTCGAAGCTAG 335

Db 388 GATTCCTGGTGGGAGCCACCATTTCTTAGCGAATGGCATTCGGAGGCGCTTCAGTGGCG 447

Qy 336 GGTTCGAAACCGCCCGCAATTCGGTTTCACGCGAGCGACGCTTTAATCAAGCCGAGACAGAT 395

Db 448 GATAGCTAAGCCACCTCAGATATGGGTACAGCAGAAAGCTCTTGTTGGGCGAGCGCAAT 507

Qy 366 TCTTCAGCTGCTTCGCGAACTCTGTCTAATCCTCGGTCTAAGAAAGATACATGAAG 455

Db 508 GCTGCAGATTGGCCCATGATCTCTCAAAACAGAGCTCCGCGACCCAGATACGACCGTGC 567

Qy 456 TCTTCTTGAATGAAGAAAGCTACAGTACATGATGTTCTTGGAATTAAGTTCCTGG 515

Db 568 GCTTTCGAGAGCCGTGATGCGGCACTCACATGAGATGTTGCTCGGAGTAAAGTTCCAGG 627

Qy 516 GGCTCTCTGCTGTATTCAGAAAGTGGTGAAGCTGAGATGATGTTCTTGCGGTTGGTGAAGC 575

Db 628 TGTGCTGTGTGTCTTCAAGAGGCTGGGGAGGCAACGTGTGTACCACTGAGAGCA 687

Qy 576 TCTGCTTAAAGAGAGGTTGCTTAAGTCTGTTAAGCAAGATGTGGTTTAAATGATGCGCT 635

Db 688 CTGTGCTTCAAGACCGTCAACCTAAGCGGTTCAAGCAGATGTGGTGTACGAATGGCATTT 747

Qy 636 TGGCTTCTCGATGTCTCGAGGATGCTATGGAATTCAGATTCACCTGATTTAATACGG 695

Db 748 GCTTATGTGACATATCAAGGAGTGTATGGCAGCAGGCCCTTCAAGATGTAATCTGCTG 807

Qy 696 TTATGATTTGTTGAGAAAGCTTTGAAGCTTTTACAGAGAGAGGAGCAAGTACCTTGC 755

Db 808 TTGTGAGGCTTTGAGAGGCACTGAGAGCTCTGCAAGAGATGGCGCAACATCTTGC 867

Qy 756 ACCGATTTACGTGCAAAATTGATGAGACTTTGGAAGAGATCACTCCGCGTATGCTT 815

Db 868 AACTGACGTCTTCAACAATTATGATGAACCTTTGAGAGAGATTACACTCGTGTGTGATTT 927

Qy 816 GGAGCTACTTGGCTTACCGCTTGATGATTTACGTGCGAAAAAGACTAAATGGTTAAG 875

Db 928 GGAGCTTCTTGCTCTCTACTGATGAAAAACATAAATTAAGCCAGAGAAAGCTTCGCA 987

Qy 876 CGGTGCGGAAATTTGTGTCGTGTTGAGAGGAGAGCAATCAGCTCTGTTGGGCG 935

Db 988 AGGTGAAAAAACAATTTGTGAGGTGTGGCAGGGGTGATTTCTACTGTGGAGAGG 1047

Qy 936 TTGACCCCGTGAAGATTTATGATGAGGCGTTTTCAGATGACAGCTGCTGAGCAGT 995

Db 1048 ATTTCCTCGAGGCGCTTATGATGAGGCTTCTTCAATGACATCAGCTGAGCAGAT 1107

Qy 996 TGAATCTTTTGAAGTACCCCAAGCAATATTCAGCAGAGTCAATTTGAAGTTACGAAGT 1055

Db 1108 GGAATTTCTTCTAATAAACCGAATAGCATACCACTGAATGTGTTGAGATCTATAGTGT 1167

Qy 1056 TGCATCTGCTCTGTGGCT 1074

Db 1168 CGCACTTGCCCAATGCT 1186

```

RESULT 12
US-10-600-070-184
; Sequence 184, Application US/10600070
; Publication No. US2004019500A1
; GENERAL INFORMATION:
; APPLICANT: Oosteryoung, Katherine W.
; APPLICANT: Vitha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of

```

```

; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 184
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Prunus persica
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-184

```

Query Match	13.7% ; Score 328.8 ; DB 7 ; Length 631 ;
Best Local Similarity	71.5% ; Pred. No. 3,4e-88 ;
Matches 451 ;	Conservative 0 ; Mismatches 168 ; Indels 12 ; Gaps 1 ;
QY	913 GGAGCATCAAGCTCTGTGTTGGGGGTTTGACCCGCGAGAAAGTTTATGAATGAGCGCTTTTAA 972
Db	1 GCAGTTGCAATTCGTGGGGGNGATTCACTACGTGMAAATTTTCATGAACGAGCGCTTCTTG 60
QY	973 CGAATGACAGTGTGAGCAGAGTTGATCTTTTGTAGCTACCCCAAGCAATATTCAGCA 1032
Db	61 CATATGACTGAGCTGAGCAGAGTTGATTTATTTGTAGCTACCCCAAGTAATATCCGGCA 120
QY	1033 GAGTCATTTGAAAGTTTACGAAAGTTCAGCTTGCTCTTGAGCTCAAGCTTTTATTTGTAG 1092
Db	121 GAAAGCTTTGAAGTTTATGGGGGTGGCTCTTGCGCTTTGTGCTCAAGCTTTGTGTGATA 180
QY	1093 AAGCCACACCTTTTACAGATGCTGATTAAGCAATTCAGACACTTCAGCAGGCTTAAGTA 1152
Db	181 AAACCTCATCACTTCAAGATGCTGAAAACCTATTCAGAACTTCAGCAGCTTAAGATA 240
QY	1153 ATGCTATGAGATTTCTCGATGTTGTATGATACAGCAATATTTGGAGATAGATTC 1212
Db	241 ACAGCTGAGGACATTTCTTTGACACTATATTAACAAAGAAAGCAGTGAATGACTTT 300
QY	1213 GGTCTAGAAAGGGGACTCTGTGCACTGTGCTTATAGGCAAAAGTTGATGAATCCGTATGTGG 1272
Db	301 GCTTTGGAGAGGGGACTCTGTTCCTCTTCTTACGGGAGCCTTGTATGACAGTGTGGG 360
QY	1273 TTGGGCTTGAACAGTGGAGATTCAATATATAGAAATCAGCTATTTGAGATTTGTTTG 1332
Db	361 TTGGGCTTGAACAGTGAATGATTCACCATATGAATTCATCTGTGTGAGCTTTGTCTG 420
QY	1333 GAGAAATTCAAATCGTGATG-----ACAAATGATGATCTTCCTGAGCTATGCAAA 1380
Db	421 GAGAACTCAAAAGATGACGATGACAAATGACAAATGACAAATGATCTTCCTGAGCTTTGAGAG 480
QY	1381 TTGTTGAAACCTGTTGGGAGAGGGGTGTCTTCTCTAGGTTACAGACACCAAGATATA 1440
Db	481 CTATTTGAGAGCTGTGTGATGAGAGTGTATTTCCCAAGTTTATAGACACCAAGACATA 540
QY	1441 AAATTTAACTCGGGGACTACTATGATGATCTTAATGTTTTGAGTTACTTGGAAAAGTG 1500
Db	541 GAGTTTCAAGCTGGGAGACTACTATGATGATCTTCAAGCTCTTGAGATACTTGAAGAGCTG 600
QY	1501 GAGGTAGTTCAAGGTTCTCTCTTTAGCTGCTG 1531
Db	601 GATGGCACTAATGTTTCAACCCCTTACGCTGCTG 631

RESULT 13
US-10-600-070-135
; Sequence 135, Application US/106000707
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseleyoung, Katherine W.
; APPLICANT: Vitka, Stanislaw
; APPLICANT: Koshanarova, Olga A.

```

; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
; US-10-600-070-135

```

```

Query Match      11.6%; Score 278.8; DB 7; Length 660;
Best Local Similarity 67.3%; Pred. No. 5.2e-73;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;

```

```

QY 81 ACCTAGCAGCAACACCTCTACAACTATCTGCTCCGAGCAAAATGGCCGACCTCTTCT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 ACCTAACCGTCTCATTTCTCCGCCCTCTCCGACCAAGTAATGGCCGAGGACTCAT 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 CTCGCACTTCAATTCACCTCGGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 TTCGCACTTCAATTCCTCGGAGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 201 CGGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 AGTCACTCTC-----ACTCTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 CATTGATTTCTACAGGATTTAGAGAGCTCAACACATTTCTTAACGATGAATTCAGAA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 TCTGCACTCTACAAATCTCTCGGCCGCAAGCAATTTCTCGGATGATTTCTGAG 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 AGCATTCGAAGCTAAGGTTTCGAAACCGCCGCAATTCGATTCGATTCGATTCGATTC 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 AGCTTATGAAGCAAAATTCGAAAGCTCTCTCAATATGCTTTCAGTAATGAGCTTTGAT 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 CAGCCGAGACAAATTTCTCAAGCTCTCTGCGAACTCTCTCTCTCTCTCTCTCTCTCTCT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 TATCTCTCTCAATTCCTCAAGCTCTCTGTAAGAACCTTATGATCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 AGAGTCAATGAAGGCTCTCTGATGATGA-----GAGCTACAGTCACTCAC 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 AGGATTAATCAAGCTCTCTCAAGTGAAGCAAGAAATGAGAAATCTTCACTTCTCAC 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 TGATGTTCTTGGGATTAAGGTTCTCTGAGGCTCTCTGATTAATGCAAGAGGATGAGAC 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 TGAATCCCTTTGACAAAGTCTCTGAGCTCTGCTGCTTTCGAGAGAGCTGAGAGAC 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 TGAGATAGTTCTTGGGTTGTAAGGCTCTGCTTAAGAGAGGTTGCTTAAGTCTTAA 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 GGAATGAGTCTTGGGATTAAGGAGGAGGTTACTGAGAGAGGTTACCGAAGATGTTAA 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 GCAAGATGAGTTTATGATTAAGGCTCTGAGGTTCTCTGATGCTCAAGGATGATGAG 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 GCAAGATGAGTTTATGATTAAGGCTCTGAGGTTCTCTGATGCTCAAGGATGATGAG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 ATTGATCCACGATTTTATTAAGGTTATGATGATTTGTTGA 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 611 TTTGTCCCGCCGAGATTTCTATTTGCTTGTGAGAGCTGGA 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

; US-10-424-599-35059
; Sequence 35059, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 35059
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_131660C.1
; US-10-424-599-35059

```

```

Query Match      9.6%; Score 231.6; DB 7; Length 1039;
Best Local Similarity 64.4%; Pred. No. 1.5e-58;
Matches 380; Conservative 0; Mismatches 204; Indels 6; Gaps 2;

```

```

QY 1823 TTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTTGAC 1882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 50 TTATGAAGAAATTAACATGCAAGTGTGCAATGATGATGATGATGATGATGATGATGAT 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1883 TGATTTCACTGTTACCCAGAGATATTTCTTAAAGCA---GCTATCTTTCAAGCA 1939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 TGGTAACTTTGTTGTTTAAATTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1940 AGGATATGTTTCTTCTATGAAATCTGATGTGCTCAATAGGCTCAGTCAAGCTGAC 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170 AATGACTGTTCTGCAATGATGTGATATCTATCAATTAATGTTCTTCTGAGATGAG 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2000 ATTCAAGAG---CACTTCCCAAGATGATCTTGAAGCTGAGCAAGAAATATATTCAGAT 2056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 AAAATGATAGAGAGCTTACCAAAATGATGATGATGATGATGATGATGATGATGATGAT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2057 GGCAGAAATTAAGTCTCTGCTTTTGGGCTGATGATCAAGCATGAAATGTTTCAAGAG 2116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 290 GGCAGAAATGTTAAATCCGAAGCTTTTGAAGCTGACCATTTGCTTGGAGAGTTGCA 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2117 TTTTGAATGGGCGAATGCTGAAGATTTGACCTGACAGAGAGCTGAACTGCGCAGCTTG 2176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 TTTTGAACGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2177 GGTGTTTATGATTAATTAACCTGTTGAACTATCTGTTGACAGTGTGACAGCTTCAGAC 2236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 410 GTTGTCTCTATGATCTACCTTGAAGAGCTTCAACATGATGATGATGATGATGATGAT 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2237 ATGAAACCCGTCTCTGTTGAAAGCACTCTGAGAGATCTGTTCTATCTGATTTG 2296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 470 ATGGGCGCGTGTGAGTGTGAAACAACTCTGAAGAGTCTACTCACTGAATGCGGTAG 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2297 TTATTCAGAAAAAATGCTACTGATGATGATGATGATGATGATGATGATGATGATGAT 2356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 530 GTCATCCGAAACAGATGCTTCTTAATGAGAGAACTTACCAACAAATATGATATGAT 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2357 GGTCAAGTCAAGGAGTGAATAATCACTGAAGGCTCTGTTCTTGATCATATA 2406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 590 TTACAGCGCGAGATGAGAAATTTGTAAGAGCTGTCTTGAAGTCTTAA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

```

; US-10-600-070-143
; Sequence 143, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oesteryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Kosharova, Olga A.
; APPLICANT: Gao Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206

```

; SOFTWARE: Patentin version 3.2
; SEQ ID NO 143
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-600-070-143

Query Match 9.3%; Score 224; DB 7; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.9e-56;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```
QY 555 AGTCTTCGGGTTGTGAGGCTCTGCTTAAGAGAGGTGCTTAAGTGGTTAAGCAGA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 AGTCTTGCAATTGAGGGCACTTACTGAGAGACCGCCCAAGCGGTTCAAGCAGA 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 615 TGTGTTTATGATGCGGCTTCCGTTCTCGATGTCTGAGGAGTCTATGCAATTGGA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 TGTGTGTGTGGCAATGGCGCTCTTAATGTGATCTATCAAGGAGCGCAATGGCGCTAG 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 675 TCCACCTGATTTTATTACTGTATGAGTTGTGAGAGAGCTTTGAACTTTTACAGA 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 CCTCCAGATGTAACTCGCTGTGTGAGTGTGAAAGGCTCTCAAGCTTTTGACAGA 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 735 GGAAGAGCAAGTAGCCTTGACCGGATTTACGTGCAAAATTGATGAGCTTTGAGA 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 GGATGGGGCAATCAATCTGCACTGGTTGTCTCTCAAAATTGATGAACCTGGAGGA 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 795 GATCACTCCGCGTTATGTCTTGAAGCTACTTGCTTACCGCTTGATGATTAAGCTGC 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 TATCACACCTCGTTGTGTGTTTGAAGCTTCTTGCCCTTCTTGATGAAAAACATCGAA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 855 GAAAAGACTAATGTGTTAAGCGGTGCGGAATATTTGTGCTGTGAGAGAGTGG 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 TGAACACCAAGAAAGTCTTGTGTGTGAGAAACATTTTGTGAGTGTGCAAGAGAG 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 915 AGCATCAGCTCTTGTGGGGGTTTGAACCGTGAGAACTTATGAAATGAGGCGTTTACG 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 TATTGTACTGTGAGAGAGATTTTGCCTGAAGCTTACATGATGAAAGCTTCTGCA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 975 AATGACAGCTGTGAGCAGGTGATCTTTTGTAGCTAACCCCAAGCAATTTCCAGCAGA 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 GATGACATCGCGAGAGATGATTTCTCTCAAAAACCGAATAGCATACCGCGCTGA 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1035 GTCAATTGAAAGTTAAGAGTGAAGTGAAGTGTGCTTGTGAGCTTCAAGCTTTTATGGA 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 ATGCTTGAATCTAATAGCTGGCACTTGCAAAATGTTGCTCAAGCAATTTGAAGTA 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 21, 2006, 15:09:20
Job time : 2718 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 14:04:13 ; Search time 545 Seconds
(without alignments)
7847.364 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagctctgacgacgt.....gtctgtcttcgcatcataa 2406

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1 COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5 COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.6	2.4	7218	2 US-08-232-463-14	Sequence 14, Appl
2	50.2	2.1	57280	3 US-09-949-016-11796	Sequence 11796, A
3	50.2	2.1	57280	3 US-09-949-016-12843	Sequence 12843, A
4	50.2	2.1	57280	3 US-09-949-016-12844	Sequence 12844, A
5	50.2	2.1	57280	3 US-09-949-016-12846	Sequence 12846, A
6	50.2	2.1	57280	3 US-09-949-016-13542	Sequence 13542, A
7	50.2	2.1	57280	3 US-09-949-016-13543	Sequence 13543, A
8	50.2	2.1	57280	3 US-09-949-016-13544	Sequence 13544, A
9	50.2	2.1	57280	3 US-09-949-016-13545	Sequence 13545, A
10	50.2	2.1	57280	3 US-09-949-016-14633	Sequence 14633, A
11	50.2	2.1	57280	3 US-09-949-016-14634	Sequence 14634, A
12	50.2	2.1	57280	3 US-09-949-016-14635	Sequence 14635, A
13	50.2	2.1	57280	3 US-09-949-016-14636	Sequence 14636, A
14	50.2	2.1	57280	3 US-09-949-016-14637	Sequence 14637, A
15	50.2	2.1	57280	3 US-09-949-016-14638	Sequence 14638, A
16	50.2	2.1	57280	3 US-09-949-016-14639	Sequence 14639, A
17	50.2	2.1	57280	3 US-09-949-016-14640	Sequence 14640, A
18	49.4	2.1	913	2 US-08-217-327-3	Sequence 3, Appl1
19	49.4	2.1	913	2 US-07-885-970A-3	Sequence 3, Appl1
20	49.4	2.1	913	2 US-08-298-687A-3	Sequence 3, Appl1
21	49.4	2.1	913	2 US-08-530-797-2	Sequence 2, Appl1
22	49.4	2.1	913	2 US-08-298-829-3	Sequence 3, Appl1
23	49.4	2.1	913	2 US-08-787-335-2	Sequence 2, Appl1
24	49.4	2.1	1984	2 US-07-885-970A-25	Sequence 25, Appl1

25	49.4	2.1	1985	2 US-08-298-687A-25	Sequence 25, Appl1
26	49.4	2.1	1985	2 US-08-298-829-25	Sequence 25, Appl1
27	49.2	2.0	1334	3 US-09-547-693-234	Sequence 234, Appl
28	48.4	2.0	152132	3 US-09-949-016-13845	Sequence 13845, A
29	48.4	2.0	152145	3 US-09-949-016-12371	Sequence 12371, A
30	46.8	1.9	114793	3 US-10-148-806-3	Sequence 3, Appl1
31	44.8	1.9	832	3 US-09-621-976-2813	Sequence 2813, Ap
32	44.2	1.8	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
33	44.2	1.8	4411529	3 US-09-103-840A-1	Sequence 22, Appl1
34	44	1.8	1141	3 US-09-806-708B-22	Sequence 8166, Ap
35	43	1.8	402	3 US-10-131-827-8166	Sequence 415, App
36	43	1.8	1281	3 US-09-991-181-415	Sequence 415, App
37	43	1.8	1281	3 US-09-990-444-415	Sequence 415, App
38	43	1.8	1281	3 US-09-997-333-415	Sequence 415, App
39	43	1.8	1281	3 US-09-992-598-415	Sequence 13, Appl
40	43	1.8	3265	3 US-09-832-129-13	Sequence 13, Appl
41	42.8	1.8	60990	3 US-09-949-016-14080	Sequence 14080, A
42	42.4	1.8	39154	3 US-09-949-016-12384	Sequence 12384, A
43	42.4	1.8	39154	3 US-09-949-016-12801	Sequence 12801, A
44	42.4	1.8	39443	3 US-09-949-016-14326	Sequence 14326, A
45	42.4	1.8	39443	3 US-09-949-016-14327	Sequence 14327, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHREIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-fls


```

; SEQ ID NO 12844
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12844

```

Query Match	2.1%	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.0059;		
Matches 103;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;

Qy	85	AGCCACAAACCTTACAAACTATGTGCTCGCCAGCAAAATGGAGCCACGCTTTCTCTCC	144
Db	13101	AGCAGCAATCACCAACAACAGCTGCTCCACCAACATACACACCTTACCAATTAAGAGCTTC	130421
Qy	145	GACTTCAATTTGACCTTCGATTCCTCTCTCTCTCTTGGCCACCGCACCAACACCGCC	204
Db	13041	ACCATCAATATCCCTATATCAAGTCCACACACACGCTTTCCACATACACACACACACAC	129828
Qy	205	ACTCTCGTCTCTTCGCCAACCATTATTAATGATGTCGCCAAGCCACGTCCTCCCATCCCAT	264
Db	12981	ACAAACCAACATCACTATCATGACTTCACACATCAACATTACATCACTCCACACCAAC	129322
Qy	265	GATTCTACCA	275
Db	12921	CATCATCACGA	12911

RESULT 5
US-09-949-016-12846/c

```

1  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
2  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
3  FILE REFERENCE: CL001307
4  CURRENT APPLICATION NUMBER: US/09/949,016
5  CURRENT FILING DATE: 2000-04-14
6  PRIOR APPLICATION NUMBER: 60/241,755
7  PRIOR FILING DATE: 2000-10-20
8  PRIOR APPLICATION NUMBER: 60/237,768
9  PRIOR FILING DATE: 2000-10-03
10 PRIOR APPLICATION NUMBER: 60/231,498
11 PRIOR FILING DATE: 2000-09-08
12 NUMBER OF SEQ ID NOS: 207012
13 SOFTWARE: FastSeq for Windows Version 4.0
14 SEQ ID NO 12846
15 LENGTH: 57280
16 TYPE: DNA
17 ORGANISM: Human
18 US-09-949-016-12846

```

	Query Match	Best Local Similarity	2.1k; Score 50.2; DB 3; Length 57280;
	Query Match	103; Conservative	53.9%; Pred. No. 0.0059; Mismatches 88; Indels 0; Gaps 0;
Qy	85 AGCCACAACACTTACACTATATCTGCTCCGACGAATAGGCGCAGCGCTCTCTCC	144	
Db	13101 AGCACATATCACACCAACTGCTCTCCACCATATCACCACTTACCAATTATGAGCTCC	13042	
Qy	145 GACTTCATTTCACCTCGAATTCCTCTCTCTCCGTGGCCACGGCGCACCAACCGCC	204	
Db	13041 ACCATCAATATTCCTTATCAGTTCACACGACCGCTTACCAATCACACACACACACC	12982	
Qy	205 ACTCTCGTCTCTCTGACCAATATATGATCGTCCGAAGCGACGTCGCCATT	264	
Db	12981 ACAACGACACACTCACTATCATGACTTCCACATCAACATATACATCACTCCGACGACAC	12922	
Qy	265 GATTTCACCA	275	
Db	12921 CATCATACGA	12911	

RESULT 6
US-09-949-016-13542/C
: Sequence 13542, Application US/09949016

```

: APPLICANT: WENTER, J. Craig et al.
: TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13542
: LENGTH: 57280
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-13542

```

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;

QY	85	AGCACAAGACCTCTCAACTATATCTGTCGACAGAAATGGGCGACCGCTCTTCTCC	144
Db	13101	AGCAACATACACACACAACTGCTCTCAGCACATACCACTTTCACATTTATGAGTCC	13042
QY	145	GACTTCATTTCACCTCCGATTCCTCTCTCTCTCTTGGCCACCGCACACACCGCC	204
Db	13041	ACCATCAATATCCCTATACAGTCCACACACACGCGTTTCACATACGACGACACAC	12987
QY	205	ACTCTGCTCTCTGCGCACCATTTATGATCGTCCGAAGCGCAGCTCCCATTCGCATT	264
Db	12981	ACAACCAACCACTCACTATCATGACTTCACACATCAATTACCATCACTCCACACAAAC	12922
QY	265	GATTTCTACCA 275	
Db	12921	CATCATCACA 12911	

```

RESULT 7
US-09-949-016-13543/C
; Sequence 13543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001037
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13543
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13543

```

Query Match	2.1%;	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.0059;		

Db 13041 ACCATCAATATCCCTATACAGTCACACCAACCGTCTTACCATACACCAACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGCCCAAGCCAGCGTCCCATCCCAT 264
Db 12981 ACAACCAACCACTACTATCATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 11

US-09-949-016-14634/c
; Sequence 14634, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14634
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14634

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGGACGAAATGGCGGACCGTCTCTCTCC 144
Db 13101 AGCACCATGACGACCAACCACTGCTCCACCATGACCACTTACCATATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTATACAGTCACACCAACCGTCTTACCATCACTCCACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGCCCAAGCCAGCGTCCCATCCCAT 264
Db 12981 ACAACCAACCACTACTATCATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 12

US-09-949-016-14635/c
; Sequence 14635, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14635
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14635

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGGACGAAATGGCGGACCGTCTCTCTCC 144
Db 13101 AGCACCATGACGACCAACCACTGCTCCACCATGACCACTTACCATATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTATACAGTCACACCAACCGTCTTACCATCACTCCACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGCCCAAGCCAGCGTCCCATCCCAT 264
Db 12981 ACAACCAACCACTACTATCATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 13

US-09-949-016-14636/c
; Sequence 14636, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14636
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14636

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTCACTATCTGCTCGGACGAAATGGCGGACCGTCTCTCTCC 144
Db 13101 AGCACCATGACGACCAACCACTGCTCCACCATGACCACTTACCATATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db 13041 ACCATCAATATCCCTATACAGTCACACCAACCGTCTTACCATCACTCCACCAAC 12982
QY 205 ACTCTGCTCTCTGCGACCATCATATGATCGCCCAAGCCAGCGTCCCATCCCAT 264
Db 12981 ACAACCAACCACTACTATCATGACTTCCACCATCAACATTACATCACTCCACCAAC 12922
QY 265 GATTCTACCA 275

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 13:35:30 ; Search time 1421 Seconds

(without alignments)
11284.483 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagcctcgtacgtcagcgt.....gctctgtctctgcatacataa 2406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406	100.0	2406	12	ADJ38129 Arabidops
2	2402.8	99.9	2406	12	ADJ38135 Arabidops
3	2399.6	99.7	2406	12	ADJ38208 Arabidops
4	2399.6	99.7	2637	12	ADJ38210 Arabidops
5	2399.6	99.7	2679	13	ADJ38210 Arabidops
6	1724	71.7	3667	12	ADJ38130 Arabidops
7	1720.8	71.5	3667	12	ADJ38136 Arabidops
8	482	20.0	561	12	ADJ38212 Arabidops
9	481	20.0	2283	12	ADJ38206 Arabidops
10	328.8	13.7	631	12	ADJ38264 Arabidops
11	278.8	11.6	660	12	ADJ38215 Arabidops
12	224	9.3	537	12	ADJ38223 Arabidops
13	222.4	9.2	552	13	ADJ38223 Arabidops
14	207	8.6	545	12	ADJ38255 Arabidops
15	197.2	8.2	491	12	ADJ38254 Arabidops
16	195.8	8.1	647	12	ADJ38265 Arabidops
17	193.8	8.1	460	12	ADJ38267 Arabidops
18	190	7.9	446	12	ADJ38231 Arabidops
19	186.6	7.8	652	12	ADJ38266 Arabidops

20	178.6	7.4	607	13	ACN48781
21	175	7.3	439	12	ADP93143
22	170.4	7.1	608	12	ADJ38217
23	170.2	7.1	563	12	ADJ38258
24	169.8	7.1	622	12	ADJ38226
25	165.4	6.9	527	12	ADJ38214
26	158.6	6.6	871	12	ADJ38232
27	155.4	6.5	307	12	ADJ38218
28	154.656	6.4	161	12	ACH86330
29	153.056	6.4	509	13	ACN59803
30	153.056	6.4	2355	13	ADJ14689
31	152.256	6.3	1880	14	ADJ18179
32	151.872	6.3	366	12	ADN99165
33	151.656	6.3	1504	13	ADJ31985
34	151.264	6.3	1470	13	ADJ38260
35	151.272	6.3	1671	13	ADJ36390
36	151.272	6.3	3245	13	ADJ27315
37	151.072	6.3	3290	13	ACN39669
38	150.856	6.3	429	12	ADJ30054
39	150.856	6.3	4068	13	ADJ40896
40	150.656	6.3	923	13	ADJ49034
41	150.656	6.3	1117	13	ADJ65456
42	150.656	6.3	1389	13	ADJ44953
43	150.656	6.3	2760	13	ADJ41745
44	150.464	6.3	909	12	ADN99181
45	150.472	6.3	1356	12	ADN01090

ALIGNMENTS

RESULT 1	ADJ38129	standard; cDNA, 2406 BP.
ID	ADJ38129	
XX	ADJ38129;	
AC	ADJ38129;	
XX		
DT	06-MAY-2004	(first entry)
DE	Arabidopsis thaliana Arc6-1 cDNA Sequ1.	
KM	prokaryotic type; plastid division; Fuz2; ARC6; ARC5; Fzo; plant cell;	
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;	
KM	herbicide target; gene; ss.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	WO2004001003-A2.	
XX		
PD	31-DEC-2003.	
XX		
PF	20-JUN-2003; 2003WO-US019536.	
XX		
PR	20-JUN-2002; 2002US-0390140P.	
PR	09-AUG-2002; 2002US-0402242P.	
XX	20-JUN-2003; 2003US-00600070.	
PA	(UNMS) UNIV MICHIGAN STATE.	
PI	Osteryoung KM, Vittha S, Koksharova OA, Gao H;	
XX	WPI, 2004-082486/08.	
DR	P-PSDB; ADJ38202.	
XX		
PT	New isolated Fuz2, ARC5 and/or Fzo-like nucleic acid sequences, useful	
PT	for further characterizing plastid division in plant cells, and in	
PT	varying agronomic and horticultural characteristics of economically	
XX	important plants.	
XX	Claim 1, SEQ ID NO 1; 287bp; English.	
PS	This invention relates to novel prokaryotic type or plastid division and	
XX	related genes and proteins. In particular, the invention relates to novel	
CC		

CC Fun2 (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;

Query Match 100.0%; Score 2406; DB 12; Length 2406;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGAAGCTCTGATGTCACGTCGGCATTTGCTCTCCCATTCATATTAAGCCATTAACA 60
DB 1 ATGGAAGCTCTGATGTCACGTCGGCATTTGCTCTCCCATTCATATTAAGCCATTAACA 60
QY 61 CCGGCGACGCAAAAGCTCCGACGTAGCACAACACCTCTACAACTATCTGCTCCGACAG 120
DB 61 CCGGCGACGCAAAAGCTCCGACGTAGCACAACACCTCTACAACTATCTGCTCCGACAG 120
QY 121 AATAGGCGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
DB 121 AATAGGCGCGACCGCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
QY 181 TTGCGCACCGCCACACACACCGCACTCTGTCTCTGCGACCACTATTAATGATGCGCC 240
DB 181 TTGCGCACCGCCACACACACCGCACTCTGTCTCTGCGACCACTATTAATGATGCGCC 240
QY 241 GAAAGCGACGTCGCCATCCCATTTGATTTTACACAGATTAAGAGCTCAACAATTC 300
DB 241 GAAAGCGACGTCGCCATCCCATTTGATTTTACACAGATTAAGAGCTCAACAATTC 300
QY 301 TTACCGATGGAATCAGAAAGCAATTCGAAGCTTTCGAAACCGCGCAATTCGAT 360
DB 301 TTACCGATGGAATCAGAAAGCAATTCGAAGCTTTCGAAACCGCGCAATTCGAT 360
QY 361 TTGAGGAGACGCTTTAATCAACCGGAGACAGATTTCTCAAGCTGTTCCGAAATCTG 420
DB 361 TTGAGGAGACGCTTTAATCAACCGGAGACAGATTTCTCAAGCTGTTCCGAAATCTG 420
QY 421 TCTAATCTCGGCTGTAAGAGAGTACAATGAAGTCTTCTGATGTAAGAAAGCTACA 480
DB 421 TCTAATCTCGGCTGTAAGAGAGTACAATGAAGTCTTCTGATGTAAGAAAGCTACA 480
QY 481 GTGATCACTGATGTTCTTGGGATTAAGGTTCTGCGGCTCTCTGTGATTAAGAAAGGT 540
DB 481 GTGATCACTGATGTTCTTGGGATTAAGGTTCTGCGGCTCTCTGTGATTAAGAAAGGT 540
QY 541 GGTGAGACTGAGATAGTTCTTGGGTTGTTGAGGCTTGTCTTAAGAGAGGTTGCTTAAG 600
DB 541 GGTGAGACTGAGATAGTTCTTGGGTTGTTGAGGCTTGTCTTAAGAGAGGTTGCTTAAG 600
QY 601 TCGTTTAAGCAAGATGTTGTTTAAATGAGCGCTTGCTTCGATGTCGAGGAGAT 660
DB 601 TCGTTTAAGCAAGATGTTGTTTAAATGAGCGCTTGCTTCGATGTCGAGGAGAT 660
QY 661 GCTATGCGATGATCCACCTGATTTTAACTGATTAAGGTTGTTGAGAAAGCTTTG 720
DB 661 GCTATGCGATGATCCACCTGATTTTAACTGATTAAGGTTGTTGAGAAAGCTTTG 720
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTGACAAATTTGAT 780
DB 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTGACAAATTTGAT 780
QY 781 GAAACTTTGAGAAAGATCACTCCGCTTATGTTGAGGACTACTGCGCTTAACGCTTGT 840
DB 781 GAAACTTTGAGAAAGATCACTCCGCTTATGTTGAGGACTACTGCGCTTAACGCTTGT 840
QY 841 GATGATTAAGCTGGAAGAACTAAATGTTTAAAGCGGTGCGGAAATATTTTGTGCT 900
DB 841 GATGATTAAGCTGGAAGAACTAAATGTTTAAAGCGGTGCGGAAATATTTTGTGCT 900
```

```
QY 901 GTTGGAGAGGTGAGCATCACTCTGTTGGGGGTTTGACCCGTGAGAAATTATGAAAT 960
DB 901 GTTGGAGAGGTGAGCATCACTCTGTTGGGGGTTTGACCCGTGAGAAATTATGAAAT 960
QY 961 GAGGCGTTTAAACGAATGACAGCTGTGAGGAGTTGATCTTTTGTACTACCTACCCAAAC 1020
DB 961 GAGGCGTTTAAACGAATGACAGCTGTGAGGAGTTGATCTTTTGTACTACCTACCCAAAC 1020
QY 1021 AATATTCAGCAGAGATCAATTGAAGTTTACGAAGTTGACCTTGTGCTCAAGCT 1080
DB 1021 AATATTCAGCAGAGATCAATTGAAGTTTACGAAGTTGACCTTGTGCTCAAGCT 1080
QY 1081 TTTATTTGTAAGAGACCACTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
DB 1081 TTTATTTGTAAGAGACCACTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
QY 1141 CAGGCTAAGCTAATGCTATGAGATTCCTGCGATGTTGATGATACAGGAATAATTTGG 1200
DB 1141 CAGGCTAAGCTAATGCTATGAGATTCCTGCGATGTTGATGATACAGGAATAATTTGG 1200
QY 1201 GAGATGACTTGGCTCTAAGAAAGGACCTGTGTGACCTCTTAATAGGCAAAATTGATGAA 1260
DB 1201 GAGATGACTTGGCTCTAAGAAAGGACCTGTGTGACCTCTTAATAGGCAAAATTGATGAA 1260
QY 1261 TGCCTGATGTTGGGCTTAAACAGTGAAGATTCACAAATATAGCAATCACTATTTGTG 1320
DB 1261 TGCCTGATGTTGGGCTTAAACAGTGAAGATTCACAAATATAGCAATCACTATTTGTG 1320
QY 1321 GAGTTGTTTGAAGATTAATCAATGCTATGACATGATGATCTCCCTGACATATGCAAA 1380
DB 1321 GAGTTGTTTGAAGATTAATCAATGCTATGACATGATGATCTCCCTGACATATGCAAA 1380
QY 1381 TTGTTGGAACCTGTTGGCAGGGGTTGCTTCCAGGTTCAAGAGACCAAAAGATAAA 1440
DB 1381 TTGTTGGAACCTGTTGGCAGGGGTTGCTTCCAGGTTCAAGAGACCAAAAGATAAA 1440
QY 1441 AATTTAACTCGGGGACTATATGATGATTCATATGTTTGAATCTTGAAGAGATG 1500
DB 1441 AATTTAACTCGGGGACTATATGATGATTCATATGTTTGAATCTTGAAGAGATG 1500
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATGCAAGATTTGAGCGGAG 1560
DB 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATGCAAGATTTGAGCGGAG 1560
QY 1561 CATGTAAAGCTAAGTCTATGACAGGACCTGCAAGAAAGTTTCTCTCCGCTATACAGAT 1620
DB 1561 CATGTAAAGCTAAGTCTATGACAGGACCTGCAAGAAAGTTTCTCTCCGCTATACAGAT 1620
QY 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATAGTACTCTGTTGGT 1680
DB 1621 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATAGTACTCTGTTGGT 1680
QY 1681 AACATGTAAGCCGTATGAGTGAAGCTGTCTTATTTAGCAAGACTGTAAAGCCCTCT 1740
DB 1681 AACATGTAAGCCGTATGAGTGAAGCTGTCTTATTTAGCAAGACTGTAAAGCCCTCT 1740
QY 1741 GAAAACCTTGAACCTAATGATTAATGCAATTCAGGCTGGGGTCTCAAGAGTACGTTGAT 1800
DB 1741 GAAAACCTTGAACCTAATGATTAATGCAATTCAGGCTGGGGTCTCAAGAGTACGTTGAT 1800
QY 1801 GAAACTACTGTTGAATGTCCTGCTGATATGTTTAAAGAGGCAAGTGTGAATCTTA 1860
DB 1801 GAAACTACTGTTGAATGTCCTGCTGATATGTTTAAAGAGGCAAGTGTGAATCTTA 1860
QY 1861 GCTGCTGTTGTCGAATTTGACTGATTTCACTGTTACGCCAGAAATATTTCTTAAAGC 1920
DB 1861 GCTGCTGTTGTCGAATTTGACTGATTTCACTGTTACGCCAGAAATATTTCTTAAAGC 1920
QY 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTCTATGAAATCTGATGCTGCTACATA 1980
DB 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTCTATGAAATCTGATGCTGCTACATA 1980
```



```

Db      ||| 841 GATGATTAGCTGCGAAAGAACTAAATGATTTAAGCGGTGCGGAATTTTGTGTC 900
Qy      ||| 901 GTTGGAGAGGTGAGACATCACTCTTTGTTGGGGTTTGAACCCGAGAGATTTAATGAT 960
Db      ||| 901 GTTGGAGAGGTGAGACATCACTCTTTGTTGGGGTTTGAACCCGAGAGATTTAATGAT 960
Qy      ||| 961 GAGGCGTTTTCGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAC 1020
Db      ||| 961 GAGGCGTTTTCGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAC 1020
Qy      ||| 1021 AATATTCACAGACAGTCAATTTGAAGTTTGAAGTTGCACTTGCTTGTGCTCAAGT 1080
Db      ||| 1021 AATATTCACAGACAGTCAATTTGAAGTTTGAAGTTGCACTTGCTTGTGCTCAAGT 1080
Qy      ||| 1081 TTTATTTGGTAAGAACCACTTTTACAGATGCTGATTAAGAAATTTCCAGCACTTCAG 1140
Db      ||| 1081 TTTATTTGGTAAGAACCACTTTTACAGATGCTGATTAAGAAATTTCCAGCACTTCAG 1140
Qy      ||| 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
Db      ||| 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
Qy      ||| 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTAATGCAAAAGTTGATGA 1260
Db      ||| 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTAATGCAAAAGTTGATGA 1260
Qy      ||| 1261 TGCCTGATGTGTTGGGCTTTAGACAGTGAAGATTCACAAATATGAAATTCAGATATTGG 1320
Db      ||| 1261 TGCCTGATGTGTTGGGCTTTAGACAGTGAAGATTCACAAATATGAAATTCAGATATTGG 1320
Qy      ||| 1321 GAGTTGTTTGGAGAAATCAATGCGATGACAAATGATGATCTCCCTGGAATGATGCAAA 1380
Db      ||| 1321 GAGTTGTTTGGAGAAATCAATGCGATGACAAATGATGATCTCCCTGGAATGATGCAAA 1380
Qy      ||| 1381 TTGTGGAACCTGTTGGCAGGGGTGTCTTTCTTAGTTTCAGAGACACCAAGATTA 1440
Db      ||| 1381 TTGTGGAACCTGTTGGCAGGGGTGTCTTTCTTAGTTTCAGAGACACCAAGATTA 1440
Qy      ||| 1441 AATATTTAACTCGGGGACTATATGATCTATGATCTTATGATTTGATTTGAAAGATG 1500
Db      ||| 1441 AATATTTAACTCGGGGACTATATGATCTTATGATCTTATGATTTGATTTGAAAGATG 1500
Qy      ||| 1501 GAGGTAAGTTCAGGCTCTCCTTAGCTGCTGAGCACTATGAGCAAGATTTGAGCCGAG 1560
Db      ||| 1501 GAGGTAAGTTCAGGCTCTCCTTAGCTGCTGAGCACTATGAGCAAGATTTGAGCCGAG 1560
Qy      ||| 1561 CATGTGAAGCTAGTCTATGACAGGACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
Db      ||| 1561 CATGTGAAGCTAGTCTATGACAGGACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
Qy      ||| 1621 AGAACTCGGCTGAAACCCAGAGATGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
Db      ||| 1621 AGAACTCGGCTGAAACCCAGAGATGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
Qy      ||| 1681 AACCAATGAGCCGCTGATGAGCTGCTGCTGCTTTATGCAAACTGTAACCCCTCT 1740
Db      ||| 1681 AACCAATGAGCCGCTGATGAGCTGCTGCTGCTTTATGCAAACTGTAACCCCTCT 1740
Qy      ||| 1741 GAAAACTTTGAAACTATATGATTTATGCAATTCAGACTGCGGCTCAAGAGATGAGTGTAT 1800
Db      ||| 1741 GAAAACTTTGAAACTATATGATTTATGCAATTCAGACTGCGGCTCAAGAGATGAGTGTAT 1800
Qy      ||| 1801 GAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGGCAAGTGTGAAGATCTTA 1860
Db      ||| 1801 GAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAGGCAAGTGTGAAGATCTTA 1860
Qy      ||| 1861 GCTGCTGCTGCTGCAATTTGACATGATTTCACTGTTACAGCCAGAAATTTTCTTAAAGC 1920
Db      ||| 1861 GCTGCTGCTGCTGCAATTTGACATGATTTCACTGTTACAGCCAGAAATTTTCTTAAAGC 1920
Qy      ||| 1921 AGCTATCTTTTCAAGCAAGATGATGTTCTTCTATGAAATCTGATGCTGCTACAT 1980

```

```

Db      ||| 1921 AGCTATCTTTTCAAGCAAGATGATGTTCTTCTATGAAATCTGATGCTTACATA 1980
Qy      ||| 1981 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTCCAGAAATGATGTAAGACTGACAG 2040
Db      ||| 1981 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTCCAGAAATGATGTAAGACTGACAG 2040
Qy      ||| 2041 AATATATGATTCAGAGTGCAGAAATTAAGTCTCTGCTTTTGGGCTGATACCGCAT 2100
Db      ||| 2041 AATATATGATTCAGAGTGCAGAAATTAAGTCTCTGCTTTTGGGCTGATACCGCAT 2100
Qy      ||| 2101 GAAATGTTAACAGAGTTTGGATGGGAAATGCTGAAGATTTGACAGACAGCAGT 2160
Db      ||| 2101 GAAATGTTAACAGAGTTTGGATGGGAAATGCTGAAGATTTGACAGACAGCAGT 2160
Qy      ||| 2161 GAACTGCGCAGCTTGGGTTGTTATGATTAATACACTGTTGAACTATCTGTGACAT 2220
Db      ||| 2161 GAACTGCGCAGCTTGGGTTGTTATGATTAATACACTGTTGAACTATCTGTGACAT 2220
Qy      ||| 2221 GTGACAGTCTCAGCAGATGAAACCCGCTGCTGTGGAAGCAACTCTGAGAGATCTGCT 2280
Db      ||| 2221 GTGACAGTCTCAGCAGATGAAACCCGCTGCTGTGGAAGCAACTCTGAGAGATCTGCT 2280
Qy      ||| 2281 TGTCTATCTGATTTGGTTCATCCAGAAAATATGCTATGATGAGAACTTACACACA 2340
Db      ||| 2281 TGTCTATCTGATTTGGTTCATCCAGAAAATATGCTATGATGAGAACTTACACACA 2340
Qy      ||| 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGTGGAAGATCACTGAAGGCTCTGTTTGA 2400
Db      ||| 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGTGGAAGATCACTGAAGGCTCTGTTTGA 2400
Qy      ||| 2401 TCATTA 2406
Db      ||| 2401 TCATTA 2406

```

```

RESULT 3
ADJ38208
ID ADJ38208 standard; DNA; 2406 BP.
XX
AC ADJ38208;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plaeid divison-related Arc6 orthologue gene 2.
XX
KW prokaryotic type; plaeid division; Fuz2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Oesteryoung KW, Vitla S, Kosharova OA, Gao H;
XX
XX MPI; 2004-082486/08.
XX
XX P-PSDB; ADJ38207.
XX
PT New isolated Fuz2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plaeid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX

```

PS Disclosure; Fig 8; 287bp; English.

XX This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Penz (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 12; Length 2406;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGATGATCGGCAATGCTCTCCCATTCATATGCGGATTCACA 60
DB 1 ATGGAAGCTCTGATGATCGGCAATGCTCTCCCATTCATATGCGGATTCACA 60
QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCAACCTCTAACAATATCTGCTCCGACG 120
DB 61 CCGGCGACGACAAAGCTCCGACGTAAGCAACCTCTAACAATATCTGCTCCGACG 120
QY 121 AATATGGGCGGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCC 180
DB 121 AATATGGGCGGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCC 180
QY 181 TTGGCCACGCGACCAACCAACGCACTCTGCTCTGCGACCACTATTTGATCGGCC 240
DB 181 TTGGCCACGCGACCAACCAACGCACTCTGCTCTGCGACCACTATTTGATCGGCC 240
QY 241 GAAAGCGACGTCGCCATCCCATTTGATTTTACAGGATTTAGAGAGCTCAACATTTCC 300
DB 241 GAAAGCGACGTCGCCATCCCATTTGATTTTACAGGATTTAGAGAGCTCAACATTTCC 300
QY 301 TTAAACCGATGGAATTCAGAAAGCAATTCGAAAGCTTTCGAAACCGCGCAATTCGGT 360
DB 301 TTAAACCGATGGAATTCAGAAAGCAATTCGAAAGCTTTCGAAACCGCGCAATTCGGT 360
QY 361 TTGAGGACGACGCTTTTATCAAGCGGACGATCTCTCTCCGACCACTATTTGATCGGCC 420
DB 361 TTGAGGACGACGCTTTTATCAAGCGGACGATCTCTCTCCGACCACTATTTGATCGGCC 420
QY 421 TCTAATCTCTGATCTAGAAAGAGTACATGAAGCTTCTTGAATGATGAAGAGCTACA 480
DB 421 TCTAATCTCTGATCTAGAAAGAGTACATGAAGCTTCTTGAATGATGAAGAGCTACA 480
QY 481 GTTATCACTGATGTTCTTGGGATTAAGTTCCTGGGCTCTCTGTGATTCGAAAGGT 540
DB 481 GTTATCACTGATGTTCTTGGGATTAAGTTCCTGGGCTCTCTGTGATTCGAAAGGT 540
QY 541 GGTGAGCTGAGATGATCTTGGGATTAAGTTCCTGGGCTCTCTGTGATTCGAAAGGT 600
DB 541 GGTGAGCTGAGATGATCTTGGGATTAAGTTCCTGGGCTCTCTGTGATTCGAAAGGT 600
QY 601 TCGTTTAAAGCAAGATGTTTAACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAG 660
DB 601 TCGTTTAAAGCAAGATGTTTAACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAG 660
QY 661 GCTATGCGCATTTGATTCATCTGATTTTAACTGATTAAGCTGATTAAGCTGATTAAG 720
DB 661 GCTATGCGCATTTGATTCATCTGATTTTAACTGATTAAGCTGATTAAGCTGATTAAG 720
QY 721 AAGCTTTTAAAGCAAGATGTTTAACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAG 780
DB 721 AAGCTTTTAAAGCAAGATGTTTAACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAG 780
QY 781 GAGACTTTGAGAGATCACTCGCGTATGCTTTGAGAGCTACTTGGCTTACCGCTTGGT 840
DB 781 GAGACTTTGAGAGATCACTCGCGTATGCTTTGAGAGCTACTTGGCTTACCGCTTGGT 840

QY 841 GATGATTAGCTGCGAAAAGACTAAATGTTTAAAGCGGTGCGGAATATTTGTGGCT 900
DB 841 GATGATTAGCTGCGAAAAGACTAAATGTTTAAAGCGGTGCGGAATATTTGTGGCT 900
QY 901 GTTGAAGAGGTGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTATGAT 960
DB 901 GTTGAAGAGGTGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTATGAT 960
QY 961 GAGCGCTTTTAAAGATGACAGCTGCTGAGCAGGTTGATCTTTTGAAGTACCCCAAGC 1020
DB 961 GAGCGCTTTTAAAGATGACAGCTGCTGAGCAGGTTGATCTTTTGAAGTACCCCAAGC 1020
QY 1021 AATATTCAGCAGATCATTTGAAGTTTAAAGTTGCACTGCTGCTGCGTCAAGCT 1080
DB 1021 AATATTCAGCAGATCATTTGAAGTTTAAAGTTGCACTGCTGCTGCGTCAAGCT 1080
QY 1081 TTTATTTGTAAGAGCACCTTTTAAAGATGCTGATTAAGCAATTCAGCAATTCAG 1140
DB 1081 TTTATTTGTAAGAGCACCTTTTAAAGATGCTGATTAAGCAATTCAGCAATTCAG 1140
QY 1141 CAGGCTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
DB 1141 CAGGCTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACAGCAATTAATGG 1200
QY 1201 GAGATGACCTTGGCTGAGAAAGGGGACCTGCGACTGCTTATTAAGCAAGTTGATGAA 1260
DB 1201 GAGATGACCTTGGCTGAGAAAGGGGACCTGCGACTGCTTATTAAGCAAGTTGATGAA 1260
QY 1261 TGCGTATGATGTTGGGCTTTAGACAGTGAAGATTCACATATAGAAATCCAGTATTTG 1320
DB 1261 TGCGTATGATGTTGGGCTTTAGACAGTGAAGATTCACATATAGAAATCCAGTATTTG 1320
QY 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGACATGATGATCTCCTGGAATATGCAAA 1380
DB 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGACATGATGATCTCCTGGAATATGCAAA 1380
QY 1381 TTGTTGAAACCTGTTGGGAGGGGTTGCTTCTCTAGGTTGAGAGCACCAAGATTA 1440
DB 1381 TTGTTGAAACCTGTTGGGAGGGGTTGCTTCTCTAGGTTGAGAGCACCAAGATTA 1440
QY 1441 AATATTTAACTCGGAGACTATGATGATCTATGATTTTGAAGTTACTTGAAGAGTG 1500
DB 1441 AATATTTAACTCGGAGACTATGATGATCTATGATTTTGAAGTTACTTGAAGAGTG 1500
QY 1501 GAGGTATTCAGGCTTCTCTTAACTGCTGCTGCACTATGCAAGATGAGATGAGCCGAG 1560
DB 1501 GAGGTATTCAGGCTTCTCTTAACTGCTGCTGCACTATGCAAGATGAGATGAGCCGAG 1560
QY 1561 CATGTGAAGCTATGCTATGACAGGCACTGCAAGATTTTCTCCGCTATPACAT 1620
DB 1561 CATGTGAAGCTATGCTATGACAGGCACTGCAAGATTTTCTCCGCTATPACAT 1620
QY 1621 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATGATGATCTGTTGCT 1680
DB 1621 AGAACTCGGCTGAACCCAGAGATGCAAGAGACAGTGTATGATGATCTGTTGCT 1680
QY 1681 AACATGTAAGCCGCTGATGATGAGCTGCTGCTTATTAAGCAAGCTGTAAGACCTCT 1740
DB 1681 AACATGTAAGCCGCTGATGATGAGCTGCTGCTTATTAAGCAAGCTGTAAGACCTCT 1740
QY 1741 GAAACTTTGAACTATATGATTAATGCAATTCAGAGCTGGGCTCAAGAGTACGTTGAT 1800
DB 1741 GAAACTTTGAACTATATGATTAATGCAATTCAGAGCTGGGCTCAAGAGTACGTTGAT 1800
QY 1801 GAAACTACTGTTGAAGATGCTCGGTGCTGATATGTTAAAGAGGCAAGTGAATCTTA 1860
DB 1801 GAAACTACTGTTGAAGATGCTCGGTGCTGATATGTTAAAGAGGCAAGTGAATCTTA 1860
QY 1861 GCTGCTGCTGCTGCAATTTGACATGATTTCACTGTTCAAGCCAGAAATATTTCTTAAAGC 1920
DB 1861 GCTGCTGCTGCTGCAATTTGACATGATTTCACTGTTCAAGCCAGAAATATTTCTTAAAGC 1920

QY 1921 AGCTCATCTTTTCACGCAAGATATGTTCTTCTATGGAATCTGATGCTGCTACCA 1980
DB 1921 AGCTCATCTTTTCACGCAAGATATGTTCTTCTATGGAATCTGATGCTGCTACCA 1980
QY 1981 GGGTCAGTCAAGCTGACGATTCAAGACCTTCCAGAAATGATGCTGATGAGAG 2040
DB 1981 GGGTCAGTCAAGCTGACGATTCAAGACCTTCCAGAAATGATGCTGATGAGAG 2040
QY 2041 AATATGTAATCCAGAGTGGCAGAAATTAAGTCTGAGCTTTTGGGCTGATCAGCCGAT 2100
DB 2041 AATATGTAATCCAGAGTGGCAGAAATTAAGTCTGAGCTTTTGGGCTGATCAGCCGAT 2100
QY 2101 GAAATGTAATCCAGAGTGGTGGATGAGCGAATGCTGAAATTTGACTGACAGAGAGCT 2160
DB 2101 GAAATGTAATCCAGAGTGGTGGATGAGCGAATGCTGAAATTTGACTGACAGAGAGCT 2160
QY 2161 GAAACTGCGCAGCTGGGCTGGTTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220
DB 2161 GAAACTGCGCAGCTGGGCTGGTTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGAGAGCAACTGAGAGAGTCTGCT 2280
DB 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGAGAGCAACTGAGAGAGTCTGCT 2280
QY 2281 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTACTGATGCTGAGAACTTACACACA 2340
DB 2281 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTACTGATGCTGAGAACTTACACACA 2340
QY 2341 AGATACGAAGTTTCTGCTGTCAGAGTCAAGGCTGGAATATCTGAGGCTCTGTTCTTGCA 2400
DB 2341 AGATACGAAGTTTCTGCTGTCAGAGTCAAGGCTGGAATATCTGAGGCTCTGTTCTTGCA 2400
QY 2401 TCATTA 2406
DB 2401 TCATTA 2406

RESULT 4
ADJ38210
ID ADJ38210 standard; DNA; 2637 BP.
XX
XX ADJ38210;
XX
DT 06-MAY-2004 (first entry)
XX
DE Placid division-related Arc6 orthologue gene 3.
XX
XX prokaryotic type; placid division; Fun2; ARC6; ARC5; Fzo; plant cell;
XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX KM herbicide target; gene; de.
OS Arabidopsis thaliana.
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003MO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX PA
XX PI Oosteryoung KM, Vitsha S, Koksharova OA, Gao H;
XX DR P-PSDB; ADJ38209.
XX
XX WPI: 2004-082486/08.
XX
XX New isolated Fun2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing placid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically

PT important plants.
XX
XX Disclosure; Fig 8; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or placid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Fun2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing placid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
XX related to the invention.
XX
SQ Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
XX
Query Match 99.7%; Score 2399.6; DB 12; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGGAAGCTCTGATGTCAGCTGCGCATTTGCTCTCTCCCATTCATTAATGCGGATTAACA 60
DB 114 ATGGAAGCTCTGATGTCAGCTGCGCATTTGCTCTCTCCCATTCATTAATGCGGATTAACA 173
QY 61 CCGGCGAGCAAAAGCTCCGAGTGGCCCAACACTCTTACAATCTGCTCCGCAAC 120
DB 174 CCGGCGAGCAAAAGCTCCGAGTGGCCCAACACTCTTACAATCTGCTCCGCAAC 233
QY 121 AATGCGCGAGCAAGCTCTCTCCGATTCATTTTCACTCCGATTCCTCTCTCTCC 180
DB 234 AATGCGCGAGCAAGCTCTCTCCGATTCATTTTCACTCCGATTCCTCTCTCTCC 293
QY 181 TTGCGCACCGCAGCACACACACCGCACTGCTCTCTCTCCACCATCTATGATGCTCC 240
DB 294 TTGCGCACCGCAGCACACACCGCACTGCTCTCTCTCCACCATCTATGATGCTCC 353
QY 241 GAAAGCGAGCTGCCCATTCCTCATTTTCTTACAGATTTAGAGACTCAAAACATTTTC 300
DB 354 GAAAGCGAGCTGCCCATTCCTCATTTTCTTACAGATTTAGAGACTCAAAACATTTTC 413
QY 301 TTAAACCGATGGAATCAAGAGATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 360
DB 414 TTAAACCGATGGAATCAAGAGATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 473
QY 361 TTCAGCGAGCAGCTTTTAATCAAGCCGAGACAGATTTCTCAAGCTCTTGCAGAACTCTG 420
DB 474 TTCAGCGAGCAGCTTTTAATCAAGCCGAGACAGATTTCTCAAGCTCTTGCAGAACTCTG 533
QY 421 TCTAATCTCTGCTTGAAGAGATCAATGAAGTCTTCTGATGATGAAGAGCTTACA 480
DB 534 TCTAATCTCTGCTTGAAGAGATCAATGAAGTCTTCTGATGATGAAGAGCTTACA 593
QY 481 GTCATACAGATGCTTCTGAGGATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 540
DB 594 GTCATACAGATGCTTCTGAGGATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 653
QY 541 GGTGAGACTGAGATAGTCTTCTGAGGATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 600
DB 654 GGTGAGACTGAGATAGTCTTCTGAGGATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 713
QY 601 TCGTTTAAAGCAAGATGTTTAAATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 660
DB 714 TCGTTTAAAGCAAGATGTTTAAATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 773
QY 661 GCTATGCAATGATGATCACTGATTTTAAATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 720
DB 774 GCTATGCAATGATGATCACTGATTTTAAATTAAGTTCCTGAGGCTCTGCTGATTTAGCAAGAGGT 833
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGCACCGGATTTACGTCGCAAAATGAT 780
DB 834 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGCACCGGATTTAGGTCGCAAAATGAT 893
QY 781 GAGACTTTGGAAGATCACTCCGCTTATGCTTGAAGCTTACCTGCTTACCGCTTGGT 840

Db 894 GAGCTTTGGAAGATCATCCGCGTATGCTTTGAGCTACTTGCTTACGCTTGT 953
 Qy 841 GATGATTAACGCTGGAAAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGCT 900
 Db 954 GATGATTAACGCTGGAAAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGCT 1013
 Qy 901 GTTGAAGAGGTGAGCATGACTCTTGTGGGGTGTGACCCGTGAGAAATTAAT 960
 Db 1014 GTTGAAGAGGTGAGCATGACTCTTGTGGGGTGTGACCCGTGAGAAATTAAT 1073
 Qy 961 GAGCGCTTTTTCAGAAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGTACC 1020
 Db 1074 GAGCGCTTTTTCAGAAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGTACC 1133
 Qy 1021 AATATTCAGAGATGATTTGAATTTAGAAATTTGACATTTGCTTTGTGGCT 1080
 Db 1134 AATATTCAGAGATGATTTGAATTTAGAAATTTGACATTTGCTTTGTGGCT 1193
 Qy 1081 TTTATTTGTAAGAACCACTTTTACAGGATGCTATTAAGCAATTCAGCACTT 1140
 Db 1194 TTTATTTGTAAGAACCACTTTTACAGGATGCTATTAAGCAATTCAGCACTT 1253
 Qy 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATTAACCGAAAT 1200
 Db 1254 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATTAACCGAAAT 1313
 Qy 1201 GAGATGACTTCGCTGTAAGAAAGGGAATCTGTGCACTGCTTATAGCAAAAT 1260
 Db 1314 GAGATGACTTCGCTGTAAGAAAGGGAATCTGTGCACTGCTTATAGCAAAAT 1373
 Qy 1261 TGCGATATGTTGGGCTTGAACGTGAGATTCACAAATTAAGCAATTCAGCA 1320
 Db 1374 TGCGATATGTTGGGCTTGAACGTGAGATTCACAAATTAAGCAATTCAGCA 1433
 Qy 1321 GAGTTGTTTGGAGATTCAAATCGATGACATGATGATCTCCCTGACATATG 1380
 Db 1434 GAGTTGTTTGGAGATTCAAATCGATGACATGATGATCTCCCTGACATATG 1493
 Qy 1381 TTGTTGAAAACCTGTTGGCAAGGGGTGCTCTTCTAGTTTCAAGAACCAAA 1440
 Db 1494 TTGTTGAAAACCTGTTGGCAAGGGGTGCTCTTCTAGTTTCAAGAACCAAA 1553
 Qy 1441 AATTTTAACTCGGGGACTATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 1554 AATTTTAACTCGGGGACTATGATGATGATGATGATGATGATGATGATGAT 1613
 Qy 1501 GAGTAGTTCAAGGTTCTCTTATGCTGCTGCACTATGCAAGATTTGAGCCG 1560
 Db 1614 GAGTAGTTCAAGGTTCTCTTATGCTGCTGCTGCACTATGCAAGATTTGAGCCG 1673
 Qy 1561 CATGTGAAGCTGATGCTATGACAGCACTGCAAGATTTTCTCCCTATATACAT 1620
 Db 1674 CATGTGAAGCTGATGCTATGACAGCACTGCAAGATTTTCTCCCTATATACAT 1733
 Qy 1621 AGAACTCGGCTGAACCAAGATGTGCAAGACAGTGTGATGATGATGATGAT 1680
 Db 1734 AGAACTCGGCTGAACCAAGATGTGCAAGACAGTGTGATGATGATGATGAT 1793
 Qy 1681 AACATGTAGCGGTGATGAGGCTGCTGCTTATGAGAAAGTATTAAGCCCT 1740
 Db 1794 AACATGTAGCGGTGATGAGGCTGCTGCTTATGAGAAAGTATTAAGCCCT 1853
 Qy 1741 GAAAACTTTGAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAGATGCT 1800
 Db 1854 GAAAACTTTGAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAGATGCT 1913
 Qy 1801 GAAACTACTGTTGAATGCTCGTTGCTGATATGTTAAAGAGGCACTGGAATCT 1860
 Db 1914 GAAACTACTGTTGAATGCTCGTTGCTGATATGTTAAAGAGGCACTGGAATCT 1973
 Qy 1861 GCTGCTGTTGGCAATGAGCTGATTCAGTGTTCAGCAGAGATTTTCTTAAAGC 1920

Db 1974 GCTGCTGTTGGCAATTTGACATGATTTCACTGTTCAGCCAGAGATTTTCTTAAAGC 2033
 Qy 1921 AGCTCATCTTTTCAAGCAAGATGATGATTTCTTATGATGATGATGATGATGAT 1980
 Db 2034 AGCTCATCTTTTCAAGCAAGATGATGATTTCTTATGATGATGATGATGATGAT 2093
 Qy 1981 GGGTCACTGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGACTGACAG 2040
 Db 2094 GGGTCACTGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGACTGACAG 2153
 Qy 2041 AATATGATTCAGAGTGGCAAGAAATTAAGTCTGTGCTTTTGGGCTGATCACCG 2100
 Db 2154 AATATGATTCAGAGTGGCAAGAAATTAAGTCTGTGCTTTTGGGCTGATCACCG 2213
 Qy 2101 GAAATGTTACAGAGTTTGAATGGCGCAATGCTGAAGATTTGACTGACAGCAGCT 2160
 Db 2214 GAAATGTTACAGAGTTTGAATGGCGCAATGCTGAAGATTTGACTGACAGCAGCT 2273
 Qy 2161 GAACTGCGAGCTTGGGTTGTTATGATTAATACATGTTGAACATATCTGTGACAGT 2220
 Db 2274 GAACTGCGAGCTTGGGTTGTTATGATTAATACATGTTGAACATATCTGTGACAGT 2333
 Qy 2221 GTGACAGTCTCAGCAAGTGAACCCGTGCTGTGGTGAAGCACTGTGAGAGTCTGCT 2280
 Db 2334 GTGACAGTCTCAGCAAGTGAACCCGTGCTGTGGTGAAGCACTGTGAGAGTCTGCT 2393
 Qy 2281 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTATGATGCTGAGAACTTACACACA 2340
 Db 2394 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTATGATGCTGAGAACTTACACACA 2453
 Qy 2341 AGATGCAAGTTTCTGCTCAGAGTCAAGGTTGCAAAATCACTGAAGGCTGTGTTCTGCA 2400
 Db 2454 AGATGCAAGTTTCTGCTCAGAGTCAAGGTTGCAAAATCACTGAAGGCTGTGTTCTGCA 2513
 Qy 2401 TCATTA 2406
 Db 2514 TCATTA 2519

RESULT 5
 ADT14901
 ID ADT14901 standard; cDNA; 2679 BP.
 XX
 AC ADT14901;
 DT 13-JAN-2005 (first entry)
 XX
 DE Plant cDNA, Seq ID 227.
 XX
 KW Plant; sb; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 OS Viridiplantae.
 OS
 PN US2004216190-A1.
 PN
 PD 28-OCT-2004.
 PD
 XX 28-DEC-2003; 2003US-00739930.
 PF
 XX 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 PR
 XX (KOVA/) KOVALIC D K.
 PA
 XX Kovalic DK;
 PI
 XX WPI; 2004-757369/74.
 DR
 XX

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetic, and in particular for producing transgenic plants with improved
PT biological characteristics.

XX Claim 1; SEQ ID NO 227; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPRO at
CC seqdata.uspro.gov/sequence.html?docid=20040216190.

XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 13; Length 2679;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	ATGGAAGCTCTGATGTCAGTTCGGGATTTGGTCTCTCCCATTCATTAAGCCATTACCA	60
DB	115	ATGGAAGCTCTGATGTCAGTTCGGGATTTGGTCTCTCCCATTCATTAAGCCATTACCA	174
QY	61	CCGGCGACGACAAAGCTCCGACGTAAGCAACAACCTCTAACAATCTCTCGCAGC	120
DB	175	CCGGCGACGACAAAGCTCCGACGTAAGCAACAACCTCTAACAATCTCTCGCAGC	234
QY	121	AAATGGGCGACCGCTCTTCTCTCGACTTCAATTCACTCCGATTCTCTCTCTCC	180
DB	235	AAATGGGCGACCGCTCTTCTCTCGACTTCAATTCACTCCGATTCTCTCTCTCC	294
QY	181	TTTGGCCACCGCCACCAACCGCACTCTGTCCTGCAACATTAATGATCGGCC	240
DB	295	TTTGGCCACCGCCACCAACCGCACTCTGTCCTGCAACATTAATGATCGGCC	354
QY	241	GAAGCGACGTCGCCATCCCATTTGATTTTACACAGATTAATGAGGCTCAACATTTG	300
DB	355	GAAGCGACGTCGCCATCCCATTTGATTTTACACAGATTAATGAGGCTCAACATTTG	414
QY	301	TTAACCGATGAATCAGAAAGCATTTGAAAGTAAAGGTTTGAACCGCGCAATTGGT	360
DB	415	TTAACCGATGAATCAGAAAGCATTTGAAAGTAAAGGTTTGAACCGCGCAATTGGT	474

QY	361	TTACGCGACGACGCTTTAATCAGCCGGAGACAGATTCTTCAAGCTGCTTGCAACTCTG	420
DB	475	TTACGCGACGACGCTTTAATCAGCCGGAGACAGATTCTTCAAGCTGCTTGCAACTCTG	534
QY	421	TCTAATCTCGGCTCTGAAAGAGATTCATTAAGAGCTCTCTTGATGATGAAGAGCTAC	480
DB	535	TCTAATCTCGGCTCTGAAAGAGATTCATTAAGAGCTCTCTTGATGATGAAGAGCTAC	594
QY	481	GTCAATCACTGATGTTCTTGGGATTAAGGTTCTGCGGAGCTCTGATATTTGCAAGAGT	540
DB	595	GTCAATCACTGATGTTCTTGGGATTAAGGTTCTGCGGAGCTCTGATATTTGCAAGAGT	654
QY	541	GGTAGACATGAGATAGTTCTTGGGATTAAGGTTCTGCGGAGCTCTGATATTTGCAAG	600
DB	655	GGTAGACATGAGATAGTTCTTGGGATTAAGGTTCTGCGGAGCTCTGATATTTGCAAG	714
QY	601	TGCTTTAAGCAAGATGCTGTTTATGTAAGCGCTTGGCTTCTGATGCTTCGAGAGAT	660
DB	715	TGCTTTAAGCAAGATGCTGTTTATGTAAGCGCTTGGCTTCTGATGCTTCGAGAGAT	774
QY	661	GCTATGGCAATGGATGCTGATTTTATTAAGTTGATGATGTTGAGGAAGCTTGG	720
DB	775	GCTATGGCAATGGATGCTGATTTTATTAAGTTGATGATGTTGAGGAAGCTTGG	834
QY	721	AAGCTTTTACAGGAG	780
DB	835	AAGCTTTTACAGGAG	894
QY	781	GAGACTTTGAAAGAGATCTCGCGCTTATGCTTTGAGACTACTTGGCTTACCGCTTGG	840
DB	895	GAGACTTTGAAAGAGATCTCGCGCTTATGCTTTGAGACTACTTGGCTTACCGCTTGG	954
QY	841	GATGATTACGCTCGCAAAAGACTAAATGCTTAAAGCGGTGAGGAGATTTTGTGCTCT	900
DB	955	GATGATTACGCTCGCAAAAGACTAAATGCTTAAAGCGGTGAGGAGATTTTGTGCTCT	1014
QY	901	GTTGAGAGAGGTGAGACATCAGCTCTTGTGGGGTTTGAACCGTGAGAAATTATGAAT	960
DB	1015	GTTGAGAGAGGTGAGACATCAGCTCTTGTGGGGTTTGAACCGTGAGAAATTATGAAT	1074
QY	961	GAGCGCTTTTACGAATGACAGCTGCTGAGAGAGTGAATCTTTTGTAGTACCCCAAGC	1020
DB	1075	GAGCGCTTTTACGAATGACAGCTGCTGAGAGAGTGAATCTTTTGTAGTACCCCAAGC	1134
QY	1021	AAATTCACGACAGAGATTTGAAGTTTACGAATGTCATCTGCTGCTGAGCTCAAGCT	1080
DB	1135	AAATTCACGACAGAGATTTGAAGTTTACGAATGTCATCTGCTGCTGAGCTCAAGCT	1194
QY	1081	TTTATTTGTAAGAGACCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCACTTCAG	1140
DB	1195	TTTATTTGTAAGAGACCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCACTTCAG	1254
QY	1141	CAGGCTAAGGTAATGCTATGAGATCTCTGCGATTTGTATGATACAGGAATTAATGG	1200
DB	1255	CAGGCTAAGGTAATGCTATGAGATCTCTGCGATTTGTATGATACAGGAATTAATGG	1314
QY	1201	GAGATGACTTGGCTTGAAGAGGAGAGCTGTGCACTGTTATTAAGCAAGTTGATGA	1260
DB	1315	GAGATGACTTGGCTTGAAGAGGAGAGCTGTGCACTGTTATTAAGCAAGTTGATGA	1374
QY	1261	TGCGATGATGTTGGGCTTTAGACAGTGAAGATTCACATTAATGAATCCAGCTATTTG	1320
DB	1375	TGCGATGATGTTGGGCTTTAGACAGTGAAGATTCACATTAATGAATCCAGCTATTTG	1434
QY	1321	GAGTTTGTGTTGAGAAATTCAAATGCTGATGACATGATGATCTCCCTGACATATGCA	1380
DB	1435	GAGTTTGTGTTGAGAAATTCAAATGCTGATGACATGATGATCTCCCTGACATATGCA	1494
QY	1381	TTGTTGAAAACCTGTTTGGAGGGGTGCTTTCTTAAGTTGAGAGACCAAGATTA	1440
DB	1495	TTGTTGAAAACCTGTTTGGAGGGGTGCTTTCTTAAGTTGAGAGACCAAGATTA	1554

OY	1441	AAATTTAACTCGGGGACATCAATATATCTCATAGTTTGAAGTTACTTGGAAGAAGTG	1500
OY	1555	AAATTTAACTCGGGGACATCAATATATCTCATAGTTTGAAGTTACTTGGAAGAAGTG	1614
OY	1501	GAGGTAGTTCAGGGTCTCTTTAGCTCTGTCGACATATGCAAGGATTTGAGCCGAG	1560
Db	1615	GAGGTAGTTCAGGGTCTCTCTTTAGCTCTGTCGACATATGCAAGGATTTGAGCCGAG	1674
OY	1561	CATGTGAAAGCTAGTGTCTATGCAAGGACTGCAAGAAAGTTTTTCTTCCCGCTATACAGAT	1620
Db	1675	CATGTGAAAGCTAGTGTCTATGCAAGGACTGCAAGAAAGTTTTTCTTCCCGCTATACAGAT	1734
OY	1621	AGAACTGGGCTGAACCCAGGAGTGTGAAGACAGGTTTAGTATGATCTGTTGGT	1680
Db	1735	AGAACTGGGCTGAACCCAGGAGTGTGAAGACAGGTTTAGTATGATCTGTTGGT	1794
OY	1681	AACATGTAGGCGCTGATGATGATGAGCCTGATGCTTTATTTGCAAGAGCTGTAAAGCCCTT	1740
Db	1795	AACATGTAGGCGCTGATGATGATGAGCCTGATGCTTTATTTGCAAGAGCTGTAAAGCCCTT	1854
OY	1741	GAAGAACTTTGAACCTAATGATTATGCAATTCGAGCTGGGGTCTCAGAGATGAGCGTTGAT	1800
Db	1855	GAAGAACTTTGAACCTAATGATTATGCAATTCGAGCTGGGGTCTCAGAGATGAGCGTTGAT	1914
OY	1801	GAATATCTATGTTGAATATGTCCTGCTGCTATATGTTAAAGAGGACAGTGTAAATCTTA	1860
Db	1915	GAATATCTATGTTGAATATGTCCTGCTGCTATATGTTAAAGAGGACAGTGTAAATCTTA	1974
OY	1861	GCTGCTGGTGTGGCAATTGGAATGATTTCACTGTTCAAGCAGAAAGTATTTCTTAAAGC	1920
Db	1975	GCTGCTGGTGTGGCAATTGGAATGATTTCACTGTTCAAGCAGAAAGTATTTCTTAAAGC	2034
OY	1921	AGCTCATCTTTTCAACGCGAAGGATATGTTTTCTTCTATGGAATCTGATGTGCTAACATA	1980
Db	2035	AGCTCATCTTTTCAACGCGAAGGATATGTTTTCTTCTATGGAATCTGATGTGCTAACATA	2094
OY	1981	GGGTCAAGTCAGAGCTGACGATTCAGAAACACTTCCCAATATGATGCTAGGACTGCAAG	2040
Db	2095	GGGTCAAGTCAGAGCTGACGATTCAGAAACACTTCCCAATATGATGCTAGGACTGCAAG	2154
OY	2041	AATATAGTATCCAAAGTGGCAGAGTTTAAGTCTGCGCTTTTGGGCCGATCACCGATA	2100
Db	2155	AATATAGTATCCAAAGTGGCAGAGTTTAAGTCTGCGCTTTTGGGCCGATCACCGATA	2214
OY	2101	GAATATGTTACAGAGGTTTTGGATGGCGAAATGCTGAAGATTTTGGACTGACAGACAGCT	2160
Db	2215	GAATATGTTACAGAGGTTTTGGATGGCGAAATGCTGAAGATTTTGGACTGACAGACAGCT	2274
OY	2161	GAATCTGGCGACGCTGGGTGGTTTATGATTATACACTGTGGAACCTATCTGTTGACAGT	2220
Db	2275	GAATCTGGCGACGCTGGGTGGTTTATGATTATACACTGTGGAACCTATCTGTTGACAGT	2334
OY	2221	GTCACAGTCTCAGACGATGGAACCCGCTGCTGCTGGAAGGAACCTGAGAGGATCTGCT	2280
Db	2335	GTCACAGTCTCAGACGATGGAACCCGCTGCTGCTGGAAGGAACCTGAGAGGATCTGCT	2394
OY	2281	TGTCATCTGATTTGGTTCATCTCAGAAACATGCTATGATGTCAAGACTTACACAACA	2340
Db	2395	TGTCATCTGATTTGGTTCATCTCAGAAACATGCTATGATGTCAAGACTTACACAACA	2454
OY	2341	AGATACGAAGTTTTCTGGTCCAGTCAAGGCTGGAAGAAATCACTGGAAGGCTCTGTTCTTGCA	2400
Db	2455	AGATACGAAGTTTTCTGGTCCAGTCAAGGCTGGAAGAAATCACTGGAAGGCTCTGTTCTTGCA	2514
OY	2401	TCATTA 2406	
Db	2515	TCATTA 2520	

AC	ADJ38130;
XX	
DT	06-MAY-2004 (first entry)
XX	
XX	Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
DE	
XX	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Pzo; plant cell;
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM	herbicide target; ds.
XX	
XX	Arabidopsis thaliana.
OS	
PV	WO2004001003-A2.
PD	
XX	
XX	31-DEC-2003.
PF	
XX	20-JUN-2003; 2003WO-US019536.
PR	
XX	20-JUN-2002; 2002US-0390140P.
PR	09-AUG-2002; 2002US-0402242P.
XX	
XX	20-JUN-2003; 2003US-00600070.
PA	(UNMS) UNIV MICHIGAN STATE.
XX	
PI	Osteryoung KM, Vicha S, Kokesarova OA, Gao H;
DR	
DR	WPI; 2004-082486/08.
DR	P-PSDB; ADJ38202.
XX	
PT	New isolated Ftn2, ARCS and/or Pzo-like nucleic acid sequences, useful
PT	for further characterizing plastid division in plant cells, and in
FT	varying agronomic and horticultural characteristics of economically
XX	important plants.
PS	
XX	Claim 1; SEQ ID NO 2; 287pp; English.
CC	This invention relates to novel prokaryotic type or plastid division and
CC	related genes and proteins. In particular, the invention relates to novel
CC	Ftn2 (ARC6), ARCS and Pzo-like genes and polypeptides. The methods and
CC	compositions of the present invention are useful for further
CC	characterising plastid division in plant cells, in order to vary
CC	agronomic and horticultural characteristics of economically important
CC	plants, such as crop, ornamental and woody plants. They can also be used
CC	as herbicide targets. The present sequence is that of a DNA sequence
CC	which is related to the invention.
SQ	
SQ	Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;
	Query Match 71.7%; Score 1724; DB 12; Length 3667;
	Best Local Similarity 79.2%; Pred. No. 0;
	Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY	
DG	1 ATGAACCTCGAGTCAAGTCGGATGGTCTCTCCCCATTGCCAATTATGCGGTACCA 60
DG	481 ATGAAACTCTGAGTCAAGTCGGATGGTCTCTCCCCATTCCAATTATGCGGTACCA 540
QY	
DG	61 CCGGCAGCAGCAAAGCTCCAGCATGAGCACAAACCTCTCAACTATCTGCTCCGAGC 120
DG	541 CCGGCAGCAGCAAAGCTCCAGCATGAGCACAAACCTCTCAACTATCTGCTCCGAGC 600
QY	
DG	121 AAATGGGCGGACGGTCTTCTCTCCGACTCAATTTCACCTCCGATTCTCTCTCTCC 180
DG	601 AAATGGGCGGACGGTCTTCTCTCCGACTCAATTTCACCTCCGATTCTCTCTCTCC 660
QY	
DG	181 TTGGCCACCGCCACACACACCGGCACCTCTGTCTTTGCGACCATCTATTGAATGCGCC 240
DG	661 TTGGCCACCGCCACACACACCGGCACCTCTGTCTTTGCGACCATCTATTGAATGCGCC 720
QY	
DG	241 GAAGCGCAGCTCCCATCCCATTTGATTTTCTACACAGGATTTAGAGCTCAACAACATTTT 300
DG	721 GAAGCGCAGCTCCCATCCCATTTGATTTTCTACACAGGATTTAGAGCTCAACAACATTTT 780
QY	
DG	301 TTAACCGATGAATCAGAGAAGCATTCGAAAGTAGGGTTTTGAAAACCGCGCAATTGCGT 360

Db 781 TTAACGATGGAATGAGAAAGACATTGAAAGCTAAGGTTTCGAAACCGCGCAATTCCGGT 840
Qy 361 TTCAGGACGACGCTTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTCCGAAACTCTG 420
Db 841 TTCAGGACGACGCTTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTCCGAAACTCTG 900
Qy 421 TCTAATCTCCGCTGAGAGAGAGTAACTGAAGGCTCTTGTATGATGAAGAAGCTACA 480
Db 901 TCTAATCTCCGCTGAGAGAGAGTAACTGAAGGCTCTTGTATGATGAAGAAGCTACA 960
Qy 481 GTCATCACTGATGCTCTTGGGAT----- 504
Db 961 GTCATCACTGATGCTCTTGGGATTAAGTAAATTTTCATTTCCGAAATATAAGTTTCTTC 1020
Qy 505 -----AAGGTTCTGGGC 518
Db 1021 GTTTTAATTCATGAATGGATTAAGAAAGAACTTTATCTAGTGAAGGTTCTTGGGC 1080
Qy 519 TCTCTGTGATTCGAAAGAGTGTGAGACTGAGATAGTTCTTCCGGTTGGTAGGCTCT 578
Db 1081 TCTCTGTGATTCGAAAGAGTGTGAGACTGAGATAGTTCTTCCGGTTGGTAGGCTCT 1140
Qy 579 GCTTAAGAGAGGTTGCTTAAGTCTGTTTAAAGAAAGTGTGTTTAAATGAGCGCTGC 638
Db 1141 GCTTAAGAGAGGTTGCTTAAGTCTGTTTAAAGAAAGTGTGTTTAAATGAGCGCTGC 1200
Qy 639 GTTTCGATGCTCGAAGGATCTATGAGATTGATCCACCTGATTTTATCTGTTA 698
Db 1201 GTTTCGATGCTCGAAGGATCTATGAGATTGATCCACCTGATTTTATCTGTTA 1260
Qy 699 TGAGTTGTTAGAGAGCTTTGAAGCTTTAC----- 720
Db 1261 TGAGTTGTTAGAGAGCTTTGAAGCTTTTACAGATGTTGACTTGTGTAATTG 1320
Qy 731 -----A 731
Db 1321 ACAGAGCTTGGCTTTAAGAACTTCTTGATTTGATTTGATTTGATGCTTGTGTA 1380
Qy 732 GAGAGAAAGAGCAAGTAGACCTTGCAACCGATTTACGTGCAAAATGATGAGCTTTGGA 791
Db 1381 GAGAGAAAGAGCAAGTAGACCTTGCAACCGATTTACGTGCAAAATGATGAGCTTTGGA 1440
Qy 792 AAGATCACTCCCGGTTATGCTTGAAGTACTTGGCTTACCGCTGGTGAATGATTAAGC 851
Db 1441 AAGATCACTCCCGGTTATGCTTGAAGTACTTGGCTTACCGCTGGTGAATGATTAAGC 1500
Qy 852 TGGAAAAAGACTAAATGCTTTAAGCGGTGCGGAATATTTGTGCTGTGAGAGAG 911
Db 1501 TGGAAAAAGACTAAATGCTTTAAGCGGTGCGGAATATTTGTGCTGTGAGAGAG 1560
Qy 912 TGGAGCATCAGCTCTTGTGGGGTTTGACCCGTGAGAGTTTATGAATGAGCGCTTTT 971
Db 1561 TGGAGCATCAGCTCTTGTGGGGTTTGACCCGTGAGAGTTTATGAATGAGCGCTTTT 1620
Qy 972 ACGAATGACGCTGCTGAGC----- 991
Db 1621 ACGAATGACGCTGCTGAGCTAATCAAGTTAGATACCTTTTAAATTTCTTAACA 1680
Qy 992 -----AGTTGATCTTTTGA 1008
Db 1681 TGAATTAACCTTAGGTTCTCATTTTAAATGATGCTGTGAGAGTTGATCTTTTGTGA 1740
Qy 1009 GCTACCCCAAGCAATATTCAGACAGACTCATTTGAAGTTTACGAAGTTGACCTTCTCT 1068
Db 1741 GCTACCCCAAGCAATATTCAGACAGACTCATTTGAAGTTTACGAAGTTGACCTTCTCT 1800
Qy 1069 GTGGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1128
Db 1801 GTGGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1860
Qy 1129 CAGCACTTGACGAGCTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1188

Db 1861 CAGCACTTGACGAGCTAAGGTAATGCTATGAGATTCCTGCCATGTTGATGATACA 1920
Qy 1189 CGAAATTAATGGGAAGATAGACTTCGCTAGAAAAGGGACTGTGCACTGCTTAATAGCC 1248
Db 1921 CGAAATTAATGGGAAGATAGACTTCGCTAGAAAAGGGACTGTGCACTGCTTAATAGCC 1980
Qy 1249 AAAAGTGAATGAAATGCCGTAATGTGTTGGCTTAAAGACATGAGAGTTCACAATATAGGAAT 1308
Db 1981 AAAAGTGAATGAAATGCCGTAATGTGTTGGCTTAAAGACATGAGAGTTCACAATATAGGAAT 2040
Qy 1309 CCAAGCTATGTAAGTGTGTTTGTGAGATTAATCGATGACAAATGATGATCTCCCT 1368
Db 2041 CCAAGCTATGTAAGTGTGTTTGTGAGATTAATCGATGACAAATGATGATCTCCCT 2100
Qy 1369 GGAATATGCAAAATGTTGGAACCTGTTGGAAGGGGTTGCTTCTTCTAGTTCAAGAGC 1428
Db 2101 GGAATATGCAAAATGTTGGAACCTGTTGGAAGGGGTTGCTTCTTCTAGTTCAAGAGC 2160
Qy 1429 ACCAAAGATPAAAAATTTAACTCGGGGACTACTATGATGATCTATGTTTGAATTAC 1488
Db 2161 ACCAAAGATPAAAAATTTAACTCGGGGACTACTATGATGATCTATGTTTGAATTAC 2220
Qy 1489 TTGGAAGAGTGAAGTACTTCAAGGTTCTCCTTTAGCTGCTGCAACTATGGAAG 1548
Db 2221 TTGGAAGAGTGAAGTACTTCAAGGTTCTCCTTTAGCTGCTGCAACTATGGAAG 2280
Qy 1549 ATTGAGCCGAGCATGTAAGGCTAGTGTCTATGACAGCACTGCAAAATTTTCTCTCC 1608
Db 2281 ATTGAGCCGAGCATGTAAGGCTAGTGTCTATGACAGCACTGCAAAATTTTCTCTCC 2340
Qy 1609 CGCTATACAGATPAAAACTCGGCTGAACCAAGATGTGCAAGAGCAAGTGTTAAGTGA 1668
Db 2341 CGCTATACAGATPAAAACTCGGCTGAACCAAGATGTGCAAGAGCAAGTGTTAAGTGA 2400
Qy 1669 GATCTGTTGCTGAACAATGAGGCTGATGAGGACCTGTGCTTTATGCAAGAGCT 1728
Db 2401 GATCTGTTGCTGAACAATGAGGCTGATGAGGACCTGTGCTTTATGCAAGAGCT 2460
Qy 1729 GTAAGACCTCTGAAAACTTTGAAACTAATGATTTATGCAATTCGACTGGGCTCTGAG 1788
Db 2461 GTAAGACCTCTGAAAACTTTGAAACTAATGATTTATGCAATTCGACTGGGCTCTGAG 2520
Qy 1789 AGTAGGTTGATGAAACTACTGTTGAATGTCCGTTGCTGATATGTTAAAGAGCAAGT 1848
Db 2521 AGTAGGTTGATGAAACTACTGTTGAATGTCCGTTGCTGATATGTTAAAGAGCAAGT 2580
Qy 1849 GTGAAGATCTGAGCTGCTGTGTGCAATTTGATGATTAATCACTGTCAGCCGAAGTAT 1908
Db 2581 GTGAAGATCTGAGCTGCTGTGTGCAATTTGATGATTAATCACTGTCAGCCGAAGTAT 2640
Qy 1909 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTCTTCTAATGAACTGAT 1968
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTCTTCTAATGAACTGAT 2700
Qy 1969 GTGCTAACCA----- 1978
Db 2701 GTGCTAACCATAGATTAATGATGAATTTTCAATATCTGATGCTCAAAATA 2760
Qy 1979 ----- 1978
Db 2761 TGCTGTTTGTGAGCTAAGAACATAGTTCCCACTTAATACATGTCACAAAGTTGACC 2820
Qy 1979 ----- 1978
Db 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGCTGAATTTTGTGATCAAA 2880
Qy 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAAGAAATACGTAGATTAG 2940
Qy 1979 -----T 1979
Db 2941 AGATTGCTTAAGTGTGCTTGTGCAACTTTTCTTCTTGATTTTCTTTTCAATTT 3000

QY 1980 AGGTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGATGGATCTAGACTGACGA 2039
 |||||
 DB 3001 AGGTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGATGGATCTAGACTGACGA 3060
 |||||
 QY 2040 GAATATATATATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGCGCTGATCA 2099
 |||||
 DB 3061 GAATATATATATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGCGCTGATCA 3120
 |||||
 QY 2100 AGAATATATATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGCGCTGATCA 2113
 |||||
 DB 3121 AGAATATATATCCAGTGGCAGAAAGATTAACTCTGCGCTTTGGCGCTGATCA 3180
 |||||
 QY 2114 -----AGCTTTGGAGTGGC 2128
 |||||
 DB 3181 TGGACATGATATATAGTCTGCTGCTTGTGATTTGATTTATATAGCTTTGATGAG 3240
 |||||
 QY 2129 GAATGCTGAAGATTGGAGTGAACAGACGCTGAACTGGCAGCTTGGGTTGGTTAG 2188
 |||||
 DB 3241 GAATGCTGAAGATTGGAGTGAACAGACGCTGAACTGGCAGCTTGGGTTGGTTAG 3300
 |||||
 QY 2189 ATTATACACTGTGAAGATTATCTGTGACAGTGTGACAGTCTGACAGATGAAACCG 2248
 |||||
 DB 3301 ATTATACACTGTGAAGATTATCTGTGACAGTGTGACAGTCTGACAGATGAAACCG 3360
 |||||
 QY 2249 CTCTGCTGAGAGCACTGCTGAGAGAGTCTGCTGATCTGATCTGATCTGATCTG 2308
 |||||
 DB 3361 CTCTGCTGAGAGCACTGCTGAGAGAGTCTGCTGATCTGATCTGATCTGATCTG 3420
 |||||
 QY 2309 ACAATGCTACTGATGTGACAGAACTGACACAAAGATTCTGTGCTGCAAGTCA 2368
 |||||
 DB 3421 ACAATGCTACTGATGTGACAGAACTGACACAAAGATTCTGTGCTGCAAGTCA 3480
 |||||
 QY 2369 GGTGGAATCACTGAAAGGCTCTGCTTCTGCTGATCTGATCTGATCTGATCTG 2406
 |||||
 DB 3481 GGTGGAATCACTGAAAGGCTCTGCTTCTGCTGATCTGATCTGATCTGATCTG 3518
 |||||
 RESULT 7
 ADJ38136
 ID ADJ38136 standard; DNA; 3667 BP.
 AC ADJ38136;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Arabidopsis thaliana AtFen2 genomic DNA SeqID10.
 XX
 KW prokaryotic type; plasmid division; Fen2; ARC6; ARC5; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402342P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oesteryoung KW, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 XX
 DR P-PSDB; ADJ38203.
 XX
 PT New isolated Fen2, ARCS and/or Fzo-like nucleic acid sequences, useful
 for further characterizing plasmid division in plant cells, and in

PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Claim 1; SEQ ID NO 10; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plasmid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fen2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plasmid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a DNA sequence
 CC which is related to the invention.
 XX
 SQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;
 Query Match 71.5%; Score 1720.8; DB 12; Length 3667;
 Best Local Similarity 79.1%; Pred. No. 0;
 Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
 QY 1 ANGAGAGCTCTGAGTACGCTGCGGATTTGCTCTCCCATTTCAATTATGCGATTACCA 60
 |||||
 DB 481 ATGGAAGCTCTGAGTACGCTGCGGATTTGCTCTCCCATTTCAATTATGCGATTACCA 540
 |||||
 QY 61 CCGGCGACGACAAAGCTCCGAGCTGACACAAACCTCTACACTATCTGCTCGGCGAGC 120
 |||||
 DB 541 CCGGCGACGACAAAGCTCCGAGCTGACACAAACCTCTACACTATCTGCTCGGCGAGC 600
 |||||
 QY 121 AATGAGGCGGACCGCTCTCTCTCCGACTTCAATTGACCTCGGATTCCTCTCTCTCC 180
 |||||
 DB 601 AATGAGGCGGACCGCTCTCTCTCTCCGACTTCAATTGACCTCGGATTCCTCTCTCTCC 660
 |||||
 QY 181 TTGGCAGCGGCGACGACGCTGCTGCTCTCTGCGACATCTATATGCTGCC 240
 |||||
 DB 661 TTGGCAGCGGCGACGACGCTGCTGCTCTCTGCGACATCTATATGCTGCC 720
 |||||
 QY 241 GAAGCGCAGCTCCGCTCCCATTTGATTTCTACAGATTTAGAGCTCAACACATTTTC 300
 |||||
 DB 721 GAAGCGCAGCTCCGCTCCCATTTGATTTCTACAGATTTAGAGCTCAACACATTTTC 780
 |||||
 QY 301 TTAAACGATGATCAGAAAGAGATTCGAAGCTAGGCTTTCGAAACCGCGCAATTCGAT 360
 |||||
 DB 781 TTAAACGATGATCAGAAAGAGATTCGAAGCTAGGCTTTCGAAACCGCGCAATTCGAT 840
 |||||
 QY 361 TTCAGCGAGAGCGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTGCGAACTCTG 420
 |||||
 DB 841 TTCAGCGAGAGCGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTGCGAACTCTG 900
 |||||
 QY 421 TCTAATCTCTGCTGTAAGAGATGATCAATGAAGTCTTCTGATGATGAAGAGCTACA 480
 |||||
 DB 901 TCTAATCTCTGCTGTAAGAGATGATCAATGAAGTCTTCTGATGATGAAGAGCTACA 960
 |||||
 QY 481 GTCATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504
 |||||
 DB 961 GTCATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 |||||
 QY 505 -----AAGTTCCTGCGGC 518
 |||||
 DB 1021 GTTTTAATTTCAATGATGATGATTAAGAAAGAACTTTATCTAGTGAAGGTTCTCTG 1080
 |||||
 QY 519 TCTCTGCTGATTTGCAAGAGAGGCTGAGACTGAGATGATTTCTGCTGCTGCTGCTGCT 578
 |||||
 DB 1081 TCTCTGCTGATTTGCAAGAGAGGCTGAGACTGAGATGATTTCTGCTGCTGCTGCTGCT 1140
 |||||
 QY 579 GCTTAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
 |||||
 DB 1141 GCTTAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 |||||
 QY 639 GTTTCGATGCTCTGAGGAGATGCTATGAGCATTTGATCACTGATTTATTACTGCTTA 698
 |||||
 DB 1201 GTTTCGATGCTCTGAGGAGATGCTATGAGCATTTGATCACTGATTTATTACTGCTTA 1260
 |||||

CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX

Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 20.0%; Score 481; DB 12; Length 2283;

Best Local Similarity 54.3%; Pred. No. 6,6e-130;

Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
QY 108 CTGCTCCGCAAGCAATGGGCGACCGCTCTCTCCGACTTCAATTCCATCCGATTC 167
DB 132 CTGGGCGGAAGCGCTCTTGGCGCACTTCCATCTCCACCGCGCGCTCCGACCC 191
QY 168 CTCTCTCTCTCTCTCTGCGCACCGCACCAACCGCACTCTGCTCTCTGCGCACATC 227
DB 192 GCCGTCCCGGCCCCCGGCCCCGCGGCGCTCCGCTCCGCTCCGCTCCGCTCT 251
QY 228 TATTGATCGTCCGGAAGCGCAAGTCCCATCCCATTTGATTTCTACAGGATATTAGAGC 287
DB 252 CCGCGAGCGCGCAAGCGCTCCCTCCGCTCCAAAGTGAATTTCTACAAAGGTTCTAGGGGC 311
QY 288 TCAACACATTTCTTAAACCATGGAATCAGAAAGCATTTGAAAGCTAGGTTTCGAACC 347
DB 312 AGAGCCACATTTCTTGGCGATGCGATCAGAGGGCGTTGAGGCAAGATAGCAAGCC 371
QY 348 GCGCAATTGCGTTTTCAGCAGCAGCGCTTAACTACGCGCGGACAGATTTCTCAAGCTGC 407
DB 372 ACCGCACTATGCTTACAGCAGAGTCTCTTGTGGTCTGCGACAAATGCTGAGATTC 431
QY 408 TTGCGAACTCTGTCTAATCTCGGTCTAGAGAGATCAATGAAAGTCTTCTTGATGA 467
DB 432 CCATGACACTCTATGAAACAGAACTCCGCACTCAGTATGATCTGCGCTTCTGAGAA 491
QY 468 TGAAGAGCTACAGTCTACTGATGTTCTTGGGATTAAGTTCTTGGGCTCTGTGT 527
DB 492 CCGTGAAGAGCTCTCACTGATGATTTGCTTGGGAAAGAGAGCGCTGG----- 540
QY 528 ATTGCAAGAGGAGGTGAGCTGAGATGTTCTTGGGTTGAGGCTGCTGCTTAAAGA 587
DB 541 -----GAGGCACTTGTCTGTGCTTGTAACTGAGAAACAGTTGCTTCTGGA 584
QY 588 GAGGTGCTTAAGTCTGTTTAAAGCAAGATGTGTTTAAAGTAAAGCGCTGCTTCTGCA 647
DB 585 TGCGGCAACCAAGCGCTTCAAGCAGAGCGTGTGCTAGCATGGCTGTGCTTATGAGA 644
QY 648 TGTCTGAGGAGTGTCTATGCAATTGATCCAAGCTGATTTTAACTGCTTATGAGTTGT 707
DB 645 TCTATCAAGGAGTGTCTATGCGAGCAAGCCTCCAGATGTAATGGCTGCGAGGTGCT 704
QY 708 TGAAGAACTTTGAAGCTTTTACAGAGGAAGAGCAAGTACCTTGCACCGGATTTACG 767
DB 705 CGAGAGGCTCTCAAGCTCTTTCAGAGAAATGAGCAAGCAATCTCCACCTGATCTGCT 764
QY 768 TGACAAATTTGATGAGACTTTTGAAGAGATCACTCCGCTTATGCTTTGAGACTTCTGG 827
DB 765 TTACAGATTTGATGAATCTCTCAGAGAGATTACCTCGCTGTATTTGAGACTTCTCTC 824
QY 828 CTTACCGCTTGTGTGATGATTAAGCTGCAAAAGCTAATGTGTTTAAAGGCTGCGGAA 887
DB 825 CTTTCCATTTGACACAGAGATCAATAAGAGCGCAAGAGGCTTCAAGGTGCGAAGAA 884
QY 888 TATTTTGTGTGTGTTGAGAGAGTGAAGCATGAGCTCTTGTGGGGGTTTGAACCGCTGA 947
DB 885 CATTGTGTGAGCGCTTTCAGAGAGGATTTGCTTACCGTTGAGAGAGATTTTCTCGTGA 944
QY 948 GAAGTTATGATGAGCGCTTTTTCAGAAATGACAGCTGCTGAGAGGTTGACTTTTGT 1007
DB 945 AGCTTATGATGAGAGCTTTTTCAGAGATGACATCAATTGAAACAGATGATTTCTTTC 1004
QY 1008 AGCTACCCCAAGCAATATTCAGAGAGATCAATTGAAAGTTTACAGAGTTCATGCTCT 1067
DB 1005 AAAAACAACGAAATGATTCCTCTGATGTGTTGAAATTTTACATGATGAGACTTTCACA 1064
```

```
QY 1068 TGTGCTCAAGCTTTATTTGTTAAGAGCCACCTTTTACAGAGATGCTGATTAAGCAAT 1127
DB 1065 TGTGCTCAAGCAATTTATTAAGGCGCAATTCATCATGATGGCGATGATCTTTT 1124
QY 1128 CCAAGCACTTTCAGCAGCTTAAGGTATGAGTATGAGATTCCTGCGATGTGATGATAC 1187
DB 1125 TGAACACTTCAGAAATTTCA-----CATAGTTCTCATTA 1160
QY 1188 ACGAATTAATTTGGAGATGAGACTTCGCTTACAAAGGAGACTCTGTGACCTGTTATAG 1247
DB 1161 TGCTTATGATTAATGATAGAGACTTTCATTTGAAAGGAGATTCCTGCTATGCTAGTCG 1220
QY 1248 CAAGTTGATGAATGCGCTATGTTGGTGGCTTACAGCTGAGAGATTCACATATGAA 1307
DB 1221 AGATTTACCAAGTGCAGATGTGGCTTGAAATGATGATCTTTCACATACAGAGA 1280
QY 1308 TCCAGCTATTTGCACTTTGTTTGAAGATTCAAATC---GTGATGACAAATGATGATCT 1364
DB 1281 CCGCAAAATTTAGAAATTTATGTCACCACTTCAGATCAGTGAAGAGAAATGATCTTCT 1340
QY 1365 CCTGACTATGCAAAATTTGTGAAACCTGTTGGCAGGGGTTGCTTCTTCTAGTTTCA 1424
DB 1341 TCCAGGGCTGTGCAACTTTTGGAGCTTGTGCTTATCTTGAAGTTTCTTACAGAGCAG 1400
QY 1425 AGACACAAAGATAAAAATTTAACTCGGGAATCTATGATGATCTATGTTTGA 1484
DB 1401 AGATATCTCGGGGATCAGATTCAGATTCAGATGATCTATGATGATTCAGAAATTTAAG 1460
QY 1485 TTTCTTGAAGAGTGAAGTATGTTCAAGGTTCTCTTTAAGCTGCTGCTGCAACTATGG 1544
DB 1461 CTACCTAAGAAAGATGAGAGGTGTGTCTTCAATTTGGCTGCTGCTGCTATATGC 1520
QY 1545 AAGATTTGAGCCGAGCATGTGAAAGCTATGCTATGCAAGCAGCTCAGAAATTTTTC 1604
DB 1521 AAACCTTGTGTCTCAAGCTACAGCTGCACTTGT----- 1555
QY 1605 TTCCGCTATACAGATGAAGAACTCGGCTGAACCAAGATGTCAGAGACAGTGTTCAG 1664
DB 1556 -----CTGTGAATCAAAATGCTATTCAGAGC----- 1581
QY 1665 TGTAGATCTGTGTTGTTAACAATGTAGCCGTGATGAGTGTGATGCTTTATTCAGAG 1724
DB 1582 -----TTCAACAA 1589
QY 1725 AGCTGAAGACCTCTGAAACCTTTGAACTAATGATTAATGCAATTCAGAGTGGGCTCTC 1784
DB 1590 GGTTTTTCATGATGAAACAGTTGACAGCTCAGCAGCAAGAAATCAATTAAGATGGCC 1649
QY 1785 AGAGATGAGCTGTATGAATCTAAGTGAATGTCGTTGCTGATGATGTTAAAGAGCG 1844
DB 1650 TGGGGGATATCTTGAAATTTTGAACAGAAATGCACTGCTCATGATTCGAATATGC 1709
QY 1845 AAGTGAAGATCTTACGCTGTGTGTGCAATTGACATGATTTCACTGTTCAGCCAGAA 1904
DB 1710 CGCTTGAAGATATCTCTGCTGCGCACTGTTTGAAGCTGTGGCAGTAAATGGGCCAA 1769
QY 1905 GTATTTTCTTAAAGAGAGCTCATCTTTCAAGCAGAGATGATGTTCTTCTTATGAATC 1964
DB 1770 ATATTTGCTCTGTAAGAGGCCCTTCTCTGCTATTAAGAGTGAAGATGATCTGTGCACT 1829
QY 1965 TGAT-----GTGCTACATAGAGGTCAAGTCAAGCTGAGATTCAGAAAGCACTTCC 2015
DB 1830 TGCTTAATAGTGTGAGCTTATCTATGATCTGTGCACTAGATGAAAGATCCAGTATATTC 1889
QY 2016 CAGAAATGATCTAGAGCTGCAAGAAATATGATTCAGATGCGCAAGAAATTAAGTCT 2075
DB 1890 TAGAATGATGCGAACTGCGAAGAAATATGTTTCCAGATGCGCAAGATTAATTA 1949
QY 2076 GGTCTTTGGGCTGATCAACCGCATAGAAATGTTACAGAGAGTTTGGATGGCGAAGTCT 2135
DB 1950 GGGCTTTGGAGCAAGAACTTTGCTGATCATTTGCAAGAGGTTCTTGAATGGCAATGCT 2009
QY 2136 GAAGATTTGAGCTGACAGAGCACTGAACCTGCGAGCTTGGGTTGATTAATATAC 2195
```

```

Db      2010 AAAGTGTGACCTGACGAGAGCGAGATTGACCGCTCATGGGTGTTCTGGAGATATAC 2069
QY      2196 ACTGTTGAACCTATCTTTGACAGTGTGACAGTCTCAGAGATGGAACCCGTGTCTGCT 2255
Db      2070 ACTATCCGAGTGTACGATGTATGATCATCATCTCTCCCTAGATGTGTGACGAGACTGT 2129
QY      2256 GGAAGCAACTCTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCATCCAGAAACAATGC 2315
Db      2130 GGAGGCTACGATGTATGAGGAGCCCAACTTACTGATGTTACTAGCCGAGAAACATGA 2189
QY      2316 TACTGATGTCAAGAACTTACCAACAAGATGAGAAAGTTTCTGTGCAAG--TCAGGGTG 2372
Db      2190 TTGATATGACACAAATAACACTACCTCCGATATGAGATGGCTTCTTCAAGCTAGAGGGTG 2249
QY      2373 GAAATTCATGGAAGGCTCTGTTCTTGATCATTA 2405
Db      2250 GAAATATACGAGAGAGAGAGCTCTCAAGTGTGA 2282

```

RESULT 10

ADJ38264 standard; cDNA; 631 BP.

```

XX      ID      ADJ38264 standard; cDNA; 631 BP.
XX      AC      ADJ38264;
XX      DT      06-MAY-2004 (first entry)
XX      DE      Plastid division-related Arc6 orthologue cDNA 51.
XX      OS      prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
XX      KM      agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX      KW      herbicide target; gene; ss.

```

OS Prunus persica.
XX PN WO2004001003-A2.

XX PD 31-DEC-2003.

XX PF 20-JUN-2003; 2003WO-US019536.

XX PR 20-JUN-2002; 2002US-0390140P.

XX PR 09-AUG-2002; 2002US-0402242P.

XX PR 20-JUN-2003; 2003US-0060070.

XX PA (UNMS) UNIV MICHIGAN STATE.

XX PI Oosteryoung KM, Vicha S, Koksharova OA, Gao H;

XX DR WPI; 2004-082486/08.

XX PT New isolated Fun2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plastid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.

XX PS Disclosure; Fig 8; 287bp; English.

XX CC This invention relates to novel prokaryotic type or plastid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Fun2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plastid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is a cDNA sequence which is
XX CC related to the invention.

XX SQ Sequence 631 BP; 174 A; 123 C; 155 G; 178 T; 0 U; 1 Other;

XX Query Match

XX Best Local Similarity 13.7%; Score 328.8; DB 12; Length 631;
XX 71.5%; Pred. No. 1.8e-85;

```

Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;
QY      913 GGAACATCAGCTCTGTTGGGGGTTTGACCCGTGAGAACTTTATGAAATGAGGCGTTTGA 972
Db      1 GCAGTTGCAATTGCTGGGGGNGATTCACTACGTGAAATTTTCATGAAAGAGGCTTTCTTG 60
QY      973 GGAATGACAGCTGTGAGACAGGTTGATCTTTTGTAGCTAACCCCAAGCAATATTCAGCA 1032
Db      61 CATATGACGTCAGGTGAGAGGTTGATTTATTTGTAGCTAACCCCAAGTAATATCCGGCA 120
QY      1033 GAGTCATTGAAAGTTTACGAAGTTGACCTGCTCTTGTGGCTCAAGCTTTTATTTGTTAG 1092
Db      121 GAAAGCTTTTAAAGTTTATGAGGGTGGCTCTTGGCCGTGTCTCAAGCTTTTGTGTTAA 180
QY      1093 AAGCCACACCTTTTACAGATGCGATGCAATTAAGCAATTCAGCAACTTCAGAGGCTAAGTA 1152
Db      181 AAACCTCATATCATTTCAAGATGCTGAAACCTTATTCAGAAACTTCAGAGCTTAAGTA 240
QY      1153 ATGGCTATGAGAGATTCCTGCGATGTTGTATGATATACGGAATTAATGGAGATAGACTTC 1212
Db      241 ACAGCTGTAGACATTTCTTGAACAATATATACCAAGAAAGAGAGGTGATAGACTTT 300
QY      1213 GGTCTAGAAAGGGGACTCTGTGACTGCTTTATAGGCAAGTTGATGATCCGTATGTG 1272
Db      301 GCTTTGAGAGGGGACTCTGTTCATCTTTCTAGGGGACCTTGATGACAGTCTGCTGG 360
QY      1273 TTGGGCTTACAGTATGAGATTCACAAATATAGAAATCCAGCTATTTGGAGTTGTTTG 1332
Db      361 TTGGGCTTACAGTATGAGATTCACAAATATAGAAATCCAGCTATTTGGAGTTGTTTG 420
QY      1333 GAGAATTCAAATGCTGATG-----ACAATGATGATCTCCCTGAGCTATGCAAA 1380
Db      421 GAGAACTCAAGAGATGACGATGACATGACAAATGACAAATGATCTTCTGAGACTTTGCAAG 480
QY      1381 TTGTTGAAACCTGTTTGGGAGGGGTTGCTTTCTTCTAGTTCAAGAGACACCAAGATPAA 1440
Db      481 CTATTTGAGAGCGTGTGATGAGGTGTATTCCTCCAGGTTTGAAGACCAAGACATATA 540
QY      1441 AAATTTAACTCGGGGACTCTATGATGATCCATATGTTTGTGATTTCTTGAAGAGTG 1500
Db      541 GAGTTGAGACTGGGAGCTATCTATGATGATCTTACGATCTTGTGAGATATCTTGAAGAGCTG 600
QY      1501 GAGGTAGTTCAGGGTTCCTCTTAGGCTGTG 1531
Db      601 GATGGCACTAATGATGATCCCTTAGCTGTG 631

```

RESULT 11

ADJ38215 standard; cDNA; 660 BP.

```

XX      ID      ADJ38215 standard; cDNA; 660 BP.
XX      AC      ADJ38215;
XX      DT      06-MAY-2004 (first entry)
XX      DE      Plastid division-related Arc6 orthologue cDNA 8.
XX      OS      prokaryotic type; plastid division; Fun2; ARC6; ARCS; Fzo; plant cell;
XX      KM      agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX      KW      herbicide target; gene; ss.

```

XX OS Medicago truncatula.

XX PN WO2004001003-A2.

XX PD 31-DEC-2003.

XX PF 20-JUN-2003; 2003WO-US019536.

XX PR 20-JUN-2002; 2002US-0390140P.

XX PR 09-AUG-2002; 2002US-0402242P.

XX PR 20-JUN-2003; 2003US-0060070.

PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 XX
 PT New isolated Fuz2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fuz2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.
 XX
 SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
 XX
 Query Match 11.6%; Score 278.8; DB 12; Length 660;
 Best Local Similarity 67.3%; Pred. No. 9.5e-71;
 Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;
 XX
 QY 81 ACCTGACCAACACACCTCTACAACTATCTGCTCCGACGAAATGGGCGACGCTTCT 140
 DB 17 ACCTAACCGTCTCATCTCCCGCCCTCTCGGACACGAAATGGGCGAGGACTCAT 76
 QY 141 CTCGCACTTCAATTTACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
 DB 77 TTCGATTTCAATTCCTCGGCGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
 QY 201 CGGCACTCTGCT 260
 DB 137 AGTCACTCTC-----ACTCTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCT 190
 QY 261 CATTGATTTCTACAGGATTTAGAGCTCAACATTTCTTAACCGATGAATCAGAG 320
 DB 191 TCTCGACCTCTACAAATCTCTCGGCGGAAAGCAATTTCTGGGTATGATTTGGAG 250
 QY 321 AGCATTCGAAGCTAGGCTTTCGAAACCGCGCAATTTGGTTTACGACGACGTTTAT 380
 DB 251 AGCTTATGAAGCAAAATTCGAAAGCTCTCTCAATGCTTTCAGTATGAAGCTTTGAT 310
 QY 381 CAGCCGAGACAGATTTCTCAAGCTGCTGGAACCTGTATCTGCTGCTAGAG 440
 DB 311 TATCTGCTGCAATTTCTCAAGCTGCTTGAAGAACCTTACGATCTGCTTCTAAG 370
 QY 441 AGAGTCAATGAAGCTTCTTGTATGATGA-----GAACTACAGTCATCAC 488
 DB 371 AGAGTATTAACAAGCTCTGTCAGCATGAAGAGATGAAGAACTTCTCATTTCTAC 430
 QY 489 TGATGTTCTTGGAGTAAGGTTCTGAGGCTCTCTGTGATTTGCAAGAGGTGTGAC 548
 DB 431 TGAATCTCTTTCGACAAAGTTCTGAGGCTGTGTGCTGTGCAAGAGCTGAGAGAC 490
 QY 549 TGAGATAGTCTTCTGGGTTGTGAGGCTGTGCTTAAGAGAGGTTGCTAAGCTTAA 608
 DB 491 GAGGTGTGCTCTGAGATGAGGAGGAGGTTTGAAGAGAGGTTTACGAAAGATGTTAA 550
 QY 609 GCAAGATGTGATTTAGTATGAGGCTTGGCTTCTGATGTCTGAGAGGATGTATGAC 668
 DB 551 GCAAGATGTGTGTGCTATGGGCTTGTGATATGTTGAGCTTTCTAAGGATGTATGAC 610
 QY 669 ATTGATCACTGATTTTATCTGATGATTAAGATTTGTTGA 710
 DB 611 TTGTGCTCCCGCAGATTTTCAATGTTGTGATGATGCTGGA 652

RESULT 12
 ADJ38223
 ID ADJ38223 standard; cDNA; 537 BP.
 XX
 AC ADJ38223;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plastid division-related Arc6 orthologue cDNA 16.
 XX
 KW prokaryotic type; plastid division; Fuz2; ARC6; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; gene; ss.
 XX
 OS Triticum aestivum.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 XX
 PT New isolated Fuz2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287bp; English.
 XX
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Fuz2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.
 XX
 SQ Sequence 537 BP; 133 A; 113 C; 151 G; 140 T; 0 U; 0 Other;
 XX
 Query Match 9.3%; Score 224; DB 12; Length 537;
 Best Local Similarity 63.6%; Pred. No. 1.1e-54;
 Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
 XX
 QY 555 AGTCTTGCGGTTGGTGGAGCTCTGCTTAAGAGAGGTTGCTTAAGCTTTAAGCAGA 614
 DB 2 AGTCTTGCAATTTGAGAGGCACTTACTTGAGAGACCGCCGCCAAGGCTTCAAGCAGA 61
 QY 615 TGTGATTTAGTATGAGCGCTTGTCTGATGTCTGAGGAGATGATGCAATTGA 674
 DB 62 TGTGATGCTGCAATGAGCGCTGCTTATGATGATATCAAGGAGCGCAATGCGGCTAG 121
 QY 675 TCACTGATTTTATTACTGTTATGAGTTGTGAGGAAGCTTTGAAGCTTTTACAGA 734
 DB 122 CCTTCAGATGATATCGCTGCTGTGAGGCTTGAAGGCTCTCAAGCTTTTTCAGA 181
 QY 735 GGAAGAGCAAGTATGCTTGCAGATTTAAGTGCAGCAATTTGATGAGATTTGAGAGA 794
 DB 182 GAGTGGGCAATCAATTCGACCTGTGCTCTCAAAATTTGATGAATCTGAGAGA 241
 QY 795 GATCACTCCGCTTATGCTTGGAGCTTACCTTGGCTTACCGCTTGGTATGATTAAGCTGC 854

DB 242 TATCACACCTGTTGTTTGGAGCTTTGGCCCTTCCCTTGATGAAAAAATCAGAA 301
 QY 855 GAAAAGACTAATAGTTTAAACCGGTGCGGATATTTTGTGCTTGTGGAGAGGTGG 914
 DB 302 TGAACCAACCAAGAGCTTCTGTGTGAGAAACATTTTGTGGGTGTGCGAGAGAG 361
 QY 915 AGCATGAGCTCTTGTGGGGGTTTGAACCGGAGAAATTAATGAATAGGCGTTTACG 974
 DB 362 TATTGTAAGCTGTTGGAGAGAGATTTTGGCTGAAGCTTCAATGAATAGCTTCTGCA 421
 QY 975 AATGACAGCTGCTGAGAGAGTTGATCTTTTGTAGTACCCCAAGCAATATTCAGAG 1034
 DB 422 GATGACATGCGCGAGAGATGATTTCTTCTCAAAAACCCGAATAGCATPACCGCTGA 481
 QY 1035 GTCAATTGAAGTTTACCAAGTTGACTGCTTGTGGCTCAAGCTTTATTTGTA 1090
 DB 482 ATGTTTGAATCTATAGCGTGGCACTTGCAAAATGTGTCAAGCAATTTGAAGTA 537

RESULT 13

ACN48855/C
 ID ACN48855 standard; cDNA; 552 BP.

ACN48855;

02-DEC-2004 (first entry)

Cotton primed seed EST clone ID: LIB3825-027-06-N6-H1, SEQ:3636.

XX Cotton; plant; EST: expressed sequence tag; transgenic plant; seed;
 XX variety DP50B; library LIB3825; molecular tag; molecular marker;
 XX genetic mapping; molecular mapping; seed germination; plant growth;
 XX plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

PN US2004123340-A1.

PD 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

PI MPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its

PT fragment, useful for isolating a variety of agronomically significant

PT genes associated with plant growth, quality or yield, and as molecular

XX tags to map genes.

XX Claim 1; SEQ ID NO 3636; 34p; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucleon313B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determining whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be

CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?docid=US20040123340

XX Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 9.2%; Score 222.4; DB 13; Length 552;

Best Local Similarity 72.2%; Pred. No. 3.3e-54;

Matches 289; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 2006 AAGCACTCCAGAAATGATGCTAGGACTGCGAGAGATATATATCCAAATGCGAGAGA 2065

DB 550 AAGAAATTAACCTAGAAATGATGCAAAATTTGCAAGAGCATTTTCCAAATGCGAGAGA 491

QY 2066 TTAAGTCTGCTGCTTTTGGGCTGATCAACCGCATAGAAATGTATACAGAGTTTGGATG 2125

DB 490 TTAATCTGAGGGGTTTGGACCTGATCACCGCTTGAATTAATGCGAGAGTTCTGGATG 431

QY 2126 GCGGAATGCTGAAGATTTTGGACTGACAGAGAGCTGAACTGGCCAGCTTGGTGT 2185

DB 430 GTCMAATGTTGAAGACATGACAGATCGTGCACCCGAAATGCTCAGCTTGGTGTAT 371

QY 2186 ATGATTAATACACTGTTGAACATATCTGTTGACAGTGTGACAGTCTGACAGATGAAACC 2245

DB 370 ATGAATTAATGCTTAATGATGAAATGCAATGCAATGCAATGCAATGCAATGCAATG 311

QY 2246 GTGCTCTGTGGAAGCAACTCTGAGAGAGTCTGCTGTATCTGATGATTTGGTTCATCAG 2305

DB 310 GAGCTGTAGTGAAGCAACTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 251

QY 2306 AAAACAATGCTACTGATGTCAGAACTTACACAAACAAAGATGCAAGTTTCTGTCAGT 2365

DB 250 AGAACAATGCTCTTAATGTAACCTCTACACACAGATATGATGATGATGATGATG 191

QY 2366 CAGGCTGGAAGAAATCAGTGAAGAGCTGCTGTTCTTCATCATTA 2405

DB 190 CAGGCTGGAAGAAATCAGTGAAGAGCTGCTGTTCTTCATCATTA 151

RESULT 14

ADJ38255

ID ADJ38255 standard; cDNA; 545 BP.

ADJ38255;

06-MAY-2004 (first entry)

XX Placid division-related Arce orthologue cDNA 42.

XX prokaryotic type; placid division; Fm2; AR6; AR5; Fzo; plant cell;

XX agronomic; horticultural; crop plant; ornamental plant; woody plant;

XX herbicide target; gene; ss.

OS Gossypium arboreum.

PN WO2004001003-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003MO-US019536.
XX PF
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-00600070.
XX (UNMS) UNIV MICHIGAN STATE.
XX PA
XX PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
XX DR MPI; 2004-082486/08.
XX
XX New isolated Fen2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plasmid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX PS Disclosure; Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plasmid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Fen2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plasmid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is a CDNA sequence which is
XX CC related to the invention.
XX
SQ Sequence 545 BP; 164 A; 107 C; 120 G; 151 T; 0 U; 3 Other;
Query Match 8.6%; Score 207; DB 12; Length 545;
Best Local Similarity 72.0%; Pred. No. 1.1e-49;
Matches 270; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 2032 ACTGCAGAGAATATAGTATCAAGTGGCAGAGATTAACTCTGCTTTGGCCCTGAT 2091
DB 2 ATTGCAGAAAGCAATTGTTGGCAAGTGGCAGAACTTAATCTGAGGCTTTGGACCTGAT 61
QY 2092 CACCGCATAGAAATGTTACCAAGAGTTTGGATGGCGAAATGCTGAAGATTGGACCTGAC 2151
DB 62 CACCGCTTGATTAATGTCAGAGGTTCTGATGCTCAAAATGTTGAAGACATGGACGAT 121
QY 2152 AGAGCAGCTGAACTGCGCAGCTTGGTGGTTATGATTATACATGTTGAACCTATCT 2211
DB 122 CGTGCAGCCGAAATCGCTCAGCTTGGTGGTATATGAATTAAGTCTAAGACATGGCC 181
QY 2212 GTTGACAGTGTGACAGTCTCAGCAGATGAAACCGTGTCTGTGTGAAGCACTGTGAG 2271
DB 182 ATTGACAGTGTTCACCTTCACTAGATGGCGAGGAGCTGATGCAAGGTAAGTCTGGA 241
QY 2272 GAGTGTGCTGTCTATTTGATTTGGTTCATTCAGAAAACAATGCTATGATGTCAAGAC 2331
DB 242 GAATTCACCTGCTGATGATGTTTCATCATCCGAGAAACAATGCTTAATGTAAATCTCC 301
QY 2332 TACACAAACAAGATAGCAAGTTTCTGGTCAAGTCAAGGTGAGAAATCACTGAAGGCTCT 2391
DB 302 TACACCAAGAGATAGATGTCTTGTTCACACTCAAGGCTGAGAAATCACTGAAGGATCT 361
QY 2392 GTTCTTGATCATTA 2406
DB 362 GTCTACAAATCTTAA 376
RESULT 15
ID ADJ38254 standard; cDNA; 491 BP.
XX ADJ38254;
AC ADJ38254;
XX
XX 06-MAY-2004 (first entry)
XX

DE Plasmid division-related Arc6 orthologue cDNA 41.
XX
XX prokaryotic type; plasmid division; Fen2; ARC6; ARCS; Fzo; plant cell;
XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX KM herbicide target; gene; ss.
XX OS
XX Trilicium aestivum.
XX PN MO2004001003-A2.
XX
XX 31-DEC-2003.
XX PD
XX
XX 20-JUN-2003; 2003MO-US019536.
XX PF
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-00600070.
XX (UNMS) UNIV MICHIGAN STATE.
XX PA
XX PI Oosteryoung KM, Viltha S, Koksharova OA, Gao H;
XX DR MPI; 2004-082486/08.
XX
XX New isolated Fen2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plasmid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX PS Disclosure; Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plasmid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Fen2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plasmid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is a CDNA sequence which is
XX CC related to the invention.
XX
SQ Sequence 491 BP; 107 A; 123 C; 142 G; 111 T; 0 U; 8 Other;
Query Match 8.2%; Score 197.2; DB 12; Length 491;
Best Local Similarity 64.1%; Pred. No. 8e-47;
Matches 295; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
QY 383 GCCGAGACAGATTTCTTCAAGCTCTTGGCAACTCTGTCTATCTCTCGCTTAAGAAAG 442
DB 2 GCCGTGGGCAAAATGCAAGTTTGAACATGATATCTCTCAACAAACAGAGCTCCGGACCG 61
QY 443 AGTACATGAAAGGCTTCTTATGATGATGAAGAACTACAGTCACTGATGTTCTTGGG 502
DB 62 AGTAAACCGCGGCTCTCTGAGAGCCGTGACCGGCGCTCAACATGATGTTGCTTGGG 121
QY 503 ATTAAGTCTTGGGCTCTCTGTGTATGCAAGAAGGTGTGAGACTGAGATAGTTCTTC 562
DB 122 ACAAGGTTCCGGGTGTGATGTCCTTCAGAGAGCTGGGAGGCAACAGGCTGTG 181
QY 563 GGGTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAGTCTTAAAGCAAGATGTGTT 622
DB 182 CAATTGAGAGCACTTAAGAGAGCCGCCGCCAAGCGTTCAAGACAGATGTGTGC 241
QY 623 TAGTTATGGGCTTGGCTTCTGATGTCTGAGAGGATGCTATGATGATGATGATGATG 682
DB 242 TGGCAATGGGCTGCTGCTTATGATGATTAATCAAGAGATGAGGCTGAGCCCTCAG 301
QY 683 ATTTATTAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 742
DB 302 ATGTAATCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
QY 743 CAGTGAAGCTTGAACCGGATTTACGTGACAAATGATGATGATGATGATGATGATGATGATG 802

Db 362 CAATCAACCTTGCACTGCTCTGCTTCACAATTTGATGAAACTGTGGAGGAGATCACAC 421
Qy 803 CGGCTTATGCTTGTGAGCTACTTGCTTACCGCTTGGTGA 842
Db 422 CTCGTGTGTGTTGTGAGCTCTTGCCCTTCTTCTTGATGA 461

Search completed: February 21, 2006, 15:41:27
Job time : 1431 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2006, 14:22:16 ; Search time 693 Seconds
(without alignments)
7370.114 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagcctcgtacgtcagct.....gctcgtcttcgtacataa 2406

Scoring table:

IDENTITY_NUC

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.New:*
1: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	46	1.9	556	6	US-09-925-065A-355459
C 2	44.8	1.9	191684	12	US-11-121-086-2
C 3	43.8	1.8	365	12	US-11-043-752-3306
C 4	43.8	1.8	365	12	US-11-043-752-3309
C 5	43	1.8	1281	6	US-10-131-826A-509
C 6	42.8	1.8	479	6	US-09-925-065A-29205
C 7	42.4	1.8	169495	12	US-11-121-086-61
C 8	42.2	1.8	1080000	8	US-10-928-446A-1
C 9	42.2	1.8	1080000	8	US-10-928-446A-181
C 10	42.2	1.8	1080000	8	US-10-928-446A-183
C 11	42.2	1.8	1080000	8	US-10-928-446A-185
C 12	42.2	1.8	1080000	8	US-10-928-446A-187
C 13	42.2	1.8	1080000	8	US-10-928-446A-189
C 14	42.2	1.8	1080000	8	US-10-928-446A-191
C 15	42.2	1.8	1080000	8	US-10-928-446A-193
C 16	42.2	1.8	1080000	8	US-10-928-446A-195
C 17	42.2	1.8	1080000	8	US-10-928-446A-197
C 18	42.2	1.8	1080000	8	US-10-928-446A-199
C 19	42.2	1.8	1080000	8	US-10-928-446A-201
C 20	41.2	1.7	2720	6	US-09-925-065A-74187

C 21	41	1.7	217	12	US-11-043-752-3461	Sequence 3461, Ap
C 22	41	1.7	262	12	US-11-043-752-3458	Sequence 3458, Ap
C 23	41	1.7	2001	12	US-11-043-752-3454	Sequence 3454, Ap
C 24	41	1.7	2001	12	US-11-043-752-3455	Sequence 3455, Ap
C 25	40.4	1.7	687	6	US-09-925-065A-69053	Sequence 69053, A
C 26	40.4	1.7	7764	12	US-11-136-527-3891	Sequence 3891, Ap
C 27	40.2	1.7	1363	6	US-09-925-065A-74546	Sequence 74546, A
C 28	40.2	1.7	31028	8	US-10-829-826B-21	Sequence 21, Appl
C 29	40.2	1.7	31028	8	US-10-829-826B-22	Sequence 22, Appl
C 30	40.2	1.7	31028	8	US-10-829-826B-24	Sequence 24, Appl
C 31	40.2	1.7	31100	8	US-10-829-826B-26	Sequence 26, Appl
C 32	39.8	1.7	1450	12	US-11-140-417-3	Sequence 3912, Ap
C 33	39.8	1.7	1450	12	US-11-140-417-3	Sequence 3912, Ap
C 34	39.8	1.7	63693	8	US-10-995-561-13269	Sequence 13269, A
C 35	39.8	1.7	168516	12	US-11-121-086-3	Sequence 3, Appl
C 36	39.4	1.6	199321	12	US-11-121-086-10	Sequence 10, Appl
C 37	39.2	1.6	11462	12	US-11-140-417-1	Sequence 1, Appl
C 38	39.2	1.6	11462	12	US-11-140-417-22	Sequence 22, Appl
C 39	39.2	1.6	120096	12	US-11-121-086-24	Sequence 24, Appl
C 40	39	1.6	523	6	US-09-925-065A-602447	Sequence 602447, A
C 41	39	1.6	10129	12	US-11-044-111-21	Sequence 21, Appl
C 42	39	1.6	153376	12	US-11-121-086-5	Sequence 5, Appl
C 43	38.8	1.6	1772	12	US-11-167-856-23	Sequence 23, Appl
C 44	38.8	1.6	7402	8	US-10-750-185-27313	Sequence 27313, A
C 45	38.8	1.6	7402	8	US-10-750-623-27313	Sequence 27313, A

ALIGNMENTS

```
RESULT 1
US-09-925-065A-355459/c
? Sequence 355459, Application US/09925065A
? Publication No. US20040181048A1
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single
? TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
? FILE REFERENCE: 108827.135
? CURRENT APPLICATION NUMBER: US/09/925, 065A
? CURRENT FILING DATE: 2001-08-08
? PRIOR APPLICATION NUMBER: US 60/243, 096
? PRIOR FILING DATE: 2000-10-24
? PRIOR APPLICATION NUMBER: US 60/252, 147
? PRIOR FILING DATE: 2000-11-20
? PRIOR APPLICATION NUMBER: US 60/250, 092
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: US 60/261, 766
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/289, 846
? PRIOR FILING DATE: 2001-05-09
? NUMBER OF SEQ ID NOS: 957086
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 355459
? LENGTH: 556
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (1)..(556)
? OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-355459

Query Match      1.9%; Score 46; DB 6; Length 556;
Best Local Similarity 60.3%; Pred. No. 0.011;
Matches 76; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY      136 CTTCTTCGAGTCATTTACATTCGATTCCTCTCTCTCTCCGACCGCCACC 195
      |||||
DB      273 CTTCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 214
      |||||
QY      196 ACCACGCGACGTCGTCGTCGTCGACCATGATGATGTCGCGAAGCCAGTCGCC 255
      |||||
```



```

Query Match 1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy      76  CTCGAGCGTAGGCACAAACACTCTTCAACACTATCTGCTCCGCGAGCAAAATGAGCGCAACCGT 135
Db      886419  CTCGCCCTCTCTCTCTCTCCCCCTCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 886478

Qy      136  CTCTCTCTCGACTTCAATTTACACTCCGATCTCTCTCCGCTCCGCTTTCGACACGCGCAAC 195
Db      886479  CTCCTCTCTCGACTCTCTACTGCCCCCTCTCTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCC 886538

Qy      196  ACCACCGCACTCTGCTCTCTCTGCGACCAATATTGATGTCGCA 242
Db      886539  TCTCTTACTCTCTCTCTCTCTTCTTTCGACACACCCGCCCCCA 886585

RESULT 9
US-10-928-446A-181
; Sequence 181, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION

```

```

1  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
2  TITLE OF INVENTION: VITAL BUDDING
3  FILE REFERENCE: 0274-5785.1US
4  CURRENT APPLICATION NUMBER: US/10/928,446A
5  CURRENT FILING DATE: 2004-08-26
6  PRIOR APPLICATION NUMBER: 60/359,741
7  PRIOR FILING DATE: 2002-02-26
8  NUMBER OF SEQ ID NOS: 202
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 181
11 LENGTH: 1080000
12 TYPE: DNA
13 ORGANISM: Homo sapiens
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: (826985)..(827008)
17 FEATURE:
18 OTHER INFORMATION: full exon 1 range is 826667-827008
19 FEATURE:
20 NAME/KEY: allele
21 LOCATION: (827008)..(827008)
22 OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
23 FEATURE:
24 NAME/KEY: CDS
25 LOCATION: (843242)..(843315)
26 OTHER INFORMATION: exon
27 FEATURE:
28 NAME/KEY: CDS
29 LOCATION: (922549)..(922630)
30 OTHER INFORMATION: exon
31 FEATURE:
32 NAME/KEY: CDS
33 LOCATION: (926021)..(926059)
34 OTHER INFORMATION: exon
35 FEATURE:
36 NAME/KEY: CDS
37 LOCATION: (929123)..(929176)
38 OTHER INFORMATION: exon
39 FEATURE:
40 NAME/KEY: CDS
41 LOCATION: (993104)..(993154)
42 OTHER INFORMATION: exon
43 FEATURE:
44 NAME/KEY: CDS
45 LOCATION: (999547)..(999608)
46 OTHER INFORMATION: exon
47 FEATURE:
48 NAME/KEY: CDS
49 LOCATION: (1000354)..(1000456)
50 OTHER INFORMATION: exon
51 FEATURE:
52 NAME/KEY: CDS
53 LOCATION: (1002118)..(1002284)
54 OTHER INFORMATION: exon
55 FEATURE:
56 NAME/KEY: CDS
57 LOCATION: (1006117)..(1006249)
58 OTHER INFORMATION: exon
59 FEATURE:
60 NAME/KEY: CDS
61 LOCATION: (1007860)..(1008036)
62 OTHER INFORMATION: exon
63 FEATURE:
64 NAME/KEY: CDS
65 LOCATION: (1010940)..(1011014)
66 OTHER INFORMATION: exon
67 FEATURE:
68 NAME/KEY: CDS
69 LOCATION: (1018160)..(1018291)
70 OTHER INFORMATION: exon
71 FEATURE:
72 NAME/KEY: CDS
73 LOCATION: (1018900)..(1018919)

```

Query Match Similarity	1.8%	Score 42.2;	DB 8;	Length 1080000;
Best Local Similarity	53.3%	Mismatch No.6.5;		
Match 89; Conservative	0;	Indels 78;	Gaps 0;	
76 CTCCGACGTAGGCACAACACTTCTAGCACTTCGCCGACGAATGGCGACGTT	135			

```
Db      886419 CTCGCCCTCCTCCGCCCCCCCCTCCCTCCCTCCCTCCCTCCGCCCCCTTC 886478
QY      136 CTCTCTCGCACTTCAATTACCTCCGAATTCCTCTCTCTCTCTGGCACCGGCACC 195
Db      886479 CTCCTCCACACTCTCACTGCCCCCTCCCTTTCCACTCTCTCTCCACCTCCCC 886538
QY      196 ACCACCGCACCTGTGCTCTCTGCGCACCATCTATTGATTCGTCGGA 242
Db      886539 TCCTTTACCTCTCTCTCTTTTGACACACCCACCCCGGCCCA 886585

RESULT 10
US-10-928-446A-183
; Sequence 183, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
; TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
; TITLE OF INVENTION: VIRAL BUDDING
; FILE REFERENCE: 0274-5785.1US
; CURRENT APPLICATION NUMBER: US/10/928,446A
; PRIOR FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/359,741
; PRIOR FILING DATE: 2002-02-26
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 183
; LENGTH: 1080000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (826985)..(827017)
; FEATURE: OTHER INFORMATION: full exon 1 range is 826667-827008
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (843242)..(843315)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (922549)..(922630)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (926021)..(926059)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (929123)..(929176)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (993104)..(993154)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (999547)..(999608)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1002118)..(1002284)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1000354)..(1000456)
; FEATURE: OTHER INFORMATION: exon
; NAME/KEY: CDS
; LOCATION: (1006117)..(1006249)
; FEATURE: OTHER INFORMATION: exon
```

```

: LOCATION:(1068609)..(1068681)
: OTHER INFORMATION: exon
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1073289)..(1073388)
: FEATURE:
: OTHER INFORMATION: full exon 30 range is 1073289-1075279
US-10-928-446A-183

Query Match 1.8%; Score 42.2; DB 8; Length 108000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Oy 76 CTCGAGCTAGCCACACACCTCTTCAACTATCTGCTCCGCGAATAGGCGCACCGT 135
Db 886419 CTCGCCCTCCCTCCGCCCTCCGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 886478

Oy 136 CTTCTCTCCGACTTCATTTTCACTCCGATTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 195
Db 886479 CTCCTCTCTCCACCTCTCTACTGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCT 886538

Oy 196 ACCACCGGCACTCTGCTCTCTGCGACCACTATGATGTCGCCGA 242
Db 886539 TCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886585

RESULT 11
US-10-928-446A-185
: Sequence 185, Application US/10928446A
: Publication No. US20050277123A1
: GENERAL INFORMATION:
: APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
: TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
: FILE REFERENCE: 0274-5785.1US
: CURRENT APPLICATION NUMBER: US/10/928,446A
: CURRENT FILING DATE: 2004-08-26
: PRIOR APPLICATION NUMBER: 60/359,741
: PRIOR FILING DATE: 2002-02-26
: NUMBER OF SEQ ID NOS: 202
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 185
: LENGTH: 1080000
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (722487)..(722534)
: FEATURE:
: OTHER INFORMATION: full length exon 1 range is 722213-722534
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (843242)..(843315)
: FEATURE:
: OTHER INFORMATION: exon
: NAME/KEY: CDS
: LOCATION: (922549)..(922630)
: FEATURE:
: OTHER INFORMATION: exon
: NAME/KEY: CDS
: LOCATION: (926021)..(926059)
: FEATURE:
: OTHER INFORMATION: exon
: NAME/KEY: CDS
: LOCATION: (929123)..(929176)
: FEATURE:
: OTHER INFORMATION: exon
: NAME/KEY: CDS
: LOCATION: (993104)..(993154)
: OTHER INFORMATION: exon
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (999547)..(999608)

```

[illegible]

NAME/KEY: CDS	LOCATION: (1006117) ..(1006249)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1007860) ..(1008036)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1010940) ..(1011014)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1018160) ..(1018291)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1018800) ..(1018919)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1020028) ..(1020225)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1026659) ..(1026736)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1028113) ..(1028167)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1034316) ..(1034374)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1041390) ..(1041455)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1043121) ..(1043350)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1044688) ..(1044989)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1047519) ..(1047589)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1050296) ..(1050391)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1060368) ..(1060441)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1062468) ..(1062708)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1064561) ..(1064620)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	LOCATION: (1066207) ..(1066314)
FEATURE: OTHER INFORMATION: exon	
NAME/KEY: CDS	

```

LOCATION: (1067768) .. (1067864)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1068609) .. (1068681)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1073289) .. (1073388)
FEATURE:
OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-187

Query Match      1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Dy       76 CTCGAGCTACGCACAACACTCTTCACTATCTGTCTCCGACGAATGGCGCACCGCT 135
Db       886419 CTCCCCCTCCTCCTCCCCCTCCCCTCCTCCTCCTCCTCCTCCTCCTCCTC 886478

Dy       136 CTCCTCTCCGACTTCATTTCACCTCCGATTCCTCCTCCTCCTCCTCCGCCACGCCACC 195
Db       886479 CTCCTCTCCACCTCTCACTGCCCCCTCCTCTTCACTCCTCCTCCCTCCACCTCCCC 886538

Dy       196 ACCACCGCACTCTGCTCTCTGCACCATCTATTGATGTCGCCA 242
Db       886539 TCTTTAACTCCTCCTCTCTTTCGACACCCACCCCGCCCCCA 886585

RESULT 13
US-10-928-446A-189
Sequence 189, Application US/10928446A
Publication No. US20050277123A1
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
FILE REFERENCE: 0274-5785.1US
CURRENT APPLICATION NUMBER: US/10/928,446A
PRIORITY FILING DATE: 2004-08-26
PRIOR APPLICATION NUMBER: 60/359,741
NUMBER OF SEQ ID NOS: 202
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 189
LENGTH: 1080000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (723508) .. (723516)
FEATURE:
OTHER INFORMATION: full length exon 1 range is 723483-723516
NAME/KEY: CDS
LOCATION: (1000354) .. (1000456)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1002118) .. (1002284)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1006117) .. (1006249)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1007860) .. (1008036)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1010940) .. (1011014)

```



```
NAME/KEY: CDS
LOCATION: (1020228) .. (1020225)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1026659) .. (1026736)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1028113) .. (1028167)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1034316) .. (1034374)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1041390) .. (1041455)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1043121) .. (1043350)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1044868) .. (1044989)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1047519) .. (1047589)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1050296) .. (1050391)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1060368) .. (1060441)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1062648) .. (1062708)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1064561) .. (1064620)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1066207) .. (1066314)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1067768) .. (1067864)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1068609) .. (1068681)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1073289) .. (1073388)
OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-191
```

```
Query Match 1.8%; Score 42.2; DB 8; Length 1080000;
Best Local Similarity 53.3%; Pred. No. 6.5;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 76 CTCGAGTTCAGCAACACCTTACATCTCTCCGCGACCAATGCGCGACCGT 135
DB 886419 CTCGCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 886478
```

```
QY 136 CTCCTCCGACTTCAATTCACCTCCGATTCCTCTCTCTCTCCGCGCGAC 195
DB 886479 CTCCTCTCCACCTCCCTCCCTCCCTCCCTCTCTCCACCTCTCTCTCCACCTCCGCC 886538
QY 196 ACCACGCGCACTCTGCTCTGCGACCATCTATTGATGTCGCCGA 242
DB 886539 TCCTTACCTCTCTCTCTCTCTCTGCGACACCCCGCGCGCCA 886585
RESULT 15
US-10-928-446A-193
Sequence 193, Application US/10928446A
Publication No. US20050277123A1
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
FILE REFERENCE: 0274-5785.1US
CURRENT APPLICATION NUMBER: US/10/928,446A
PRIORITY FILING DATE: 2004-08-26
PRIORITY FILING DATE: 2002-02-26
NUMBER OF SEQ ID NOS: 202
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 193
LENGTH: 1080000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1000354) .. (1000456)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (999562) .. (999608)
OTHER INFORMATION: full length exon 7 range is 999547-999608
FEATURE:
NAME/KEY: CDS
LOCATION: (1002118) .. (1002284)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1006117) .. (1006249)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1007660) .. (1008036)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1010940) .. (1011014)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018160) .. (1018291)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1018600) .. (1018919)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1020028) .. (1020225)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1026659) .. (1026736)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1028113) .. (1028167)
```


Search completed: February 21, 2006, 14:58:05
Job time : 699 secs

```

OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1034316) .. (1034374)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1041390) .. (1041455)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1043121) .. (1043350)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1044868) .. (1044989)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1047519) .. (1047589)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1050296) .. (1050391)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1060368) .. (1060441)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1062648) .. (1062708)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1064561) .. (1064620)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1066207) .. (1066314)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1067768) .. (1067864)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1068609) .. (1068681)
OTHER INFORMATION: exon
FEATURE:
NAME/KEY: CDS
LOCATION: (1073289) .. (1073388)
FEATURE:
OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-193

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 21, 2006, 12:51:25 / Search time 10553 seconds
(without alignments)
4314.565 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGIGLSPFQCRLLP.....YEVFWKSGWKITGSLVAS 801

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=rlh
-Q=/abs/ABSWEB.spool/US10600070/runat_21022006_112358_8614/app.query.fasta-1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absg03h
-USER=US10600070 @CCN_1_1_4939 @runat_21022006_112358_8614 -NCPU=6 -ICU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV -TIMEOUT=120
-MARN -TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_da.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_str.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4052	99.7	2438	15	AY221469 Arabidops
2	4051	99.7	2436	15	AY150490 Arabidops
3	4051	99.7	2637	15	AY091075 Arabidops

Result No.	Score	Query Match	Length	DB ID	Description
4	3837.5	94.4	3668	15	AY221468 Arabidops
5	3825.5	94.2	85791	15	AB016888 Arabidops
6	3823.5	94.1	3664	15	AY221467 Arabidops
7	1540.5	37.9	110000	15	AP008208_011 Arabidops
8	1540.5	37.9	150452	15	AP004885 Arabidops
9	1241	30.5	133779	14	AC158210 Arabidops
10	1222	30.1	184752	14	AC155597 Arabidops
11	702	17.3	110000	14	CR954204 Arabidops
12	497	12.2	110000	14	BA000019_33 Arabidops
13	497	12.2	110000	14	BA000019_33 Arabidops
14	449	11.1	110000	14	BA000045_36 Arabidops
15	446	11.0	2469	14	AF421196 Arabidops
16	446	11.0	110000	14	AP008231_22 Arabidops
17	436.5	10.7	110000	14	BA000022_23 Arabidops
18	431	10.6	2857	15	AY074283 Arabidops
19	425	10.5	110000	14	BA000039_07 Arabidops
20	357.5	8.8	110000	14	CP000095_07 Arabidops
21	357.5	8.8	110000	14	CP000095_08 Arabidops
22	351	8.6	550	10	DQ022571 Arabidops
23	344.5	8.5	348986	14	BX572093 Arabidops
24	340	8.4	349213	14	BX565693 Arabidops
25	323	7.9	349323	14	BX572096 Arabidops
26	318	7.8	300143	14	AE017165 Arabidops
27	299	7.4	117818	14	AC160013 Arabidops
28	229.5	5.6	110000	15	AP008210_345 Arabidops
29	229.5	5.6	164943	15	OSJG000052 Arabidops
30	227.5	5.6	104146	14	AP008231_23 Arabidops
31	220	5.4	110000	14	AP000419 Arabidops
32	185	4.6	81875	15	AY596297_28 Arabidops
33	167.5	4.1	110000	14	AB056299 Arabidops
34	155.5	3.8	41599	6	B63260 Arabidops
35	155.5	3.8	41599	6	B63260 Arabidops
36	151.5	3.7	41884	14	BAM575642 Arabidops
37	150	3.7	11237	14	AE007907 Arabidops
38	150	3.7	12027	14	AE008959 Arabidops
39	149	3.7	110000	14	AP006878_09 Arabidops
40	148.5	3.7	6270	6	I26720 Arabidops
41	148.5	3.7	80841	6	CO975935 Arabidops
42	148.5	3.7	122883	14	BAM575417 Arabidops
43	148.5	3.7	122883	14	BAM576102 Arabidops
44	148	3.6	46513	1	AY137375 Arabidops
45	146.5	3.6	110000	1	CR555306_05 Arabidops

ALIGNMENTS

RESULT 1
LOCUS AY221469
DEFINITION Arabidopsis thaliana division protein (ARC6) mRNA, complete cds;
ACCESSION AY221469
VERSION AY221469
KEYWORDS GI:33436353
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 2438)
AUTHORS Oesteryoung,K.W.
TITLE ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL Plant Cell 15 (8), 1918-1933 (2003)
PUBMED 12897262
REFERENCE 2 (bases 1 to 2438)
AUTHORS Vitsha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Exp,H. and Oesteryoung,K.W.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES Location/Qualifiers

```

source
1. .2438
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Wassilewskija"
1. .2438
/gene="ARC6"
/ncbi_synonym="At5g42480"
2. .2407
/gene="ARC6"
/codon_start=1
/product="division protein"
/protein_id="AA018646.1"
/db_xref="gi:33436354"
/translation="MEALSHVGIGLSPPOLCRLEPPATTKLRKRNSTSTTCASAKMAD
RLDLEFNFTSSSSSPATATTTATVTLSPSIDIPIRPHVPIPIDFYQVGAQTHPLT
GIRRAFEARVSKPPOFSDDALISROIQAACETLSNPSRREYNEGILDEGAT
VITPMDKPVGALCVLOEGETETIVRGALAKERLPKSFQDVVLAALADVS
RDNALMDPDPIITNGTVEFEALKILQEGASLAPDLRAOIDETLEETTPRYELLGL
LPLEDDDAARKLNELSGVRNLINVSVGGGASALVIGITREKPNEARLPMIAEONVLL
FVATPSNIPASFEVEYVALVAQAFLGKPHLLQDADKFOQLQAATMAEIPAM
LYDRNMWEIDFGLERGICALLIGKVDCRMWLDSDSQRNPAIYERVLANSND
DNDLPGICLKLLETMLAGVPPRPDRTRDKKFKLGDIYDDPMVLISYERVENVGSP
AAATAMARIGAEHYKASMAQLOKVFPSRYDRNASBKQOETVFSVDYEVGNVGD
GEPVPIFAEVRPSENFTNDYAIRAGVSESVDETTEMVSVAEMLEKASAKILAAIV
AIGLISFSOKYPLKSSSPQRKDMVSMESSDVATITSVRDSLEALPRMDARAEKI
VSKQOKIKSLAFGPDHRIEMLPETLDGMLKIWTRDRAETQQLQGLVDYDITLKISNS
VTSADGTRALVEATLESACLSDLHPENNATVRYITRIEYFMSKSGKLTIGSV
LAS"

misc_difference
474
/gene="ARC6"
/ncbi_synonym="resulte in Gly to Glu substitution"
/replace="a"
misc_difference
1831
/gene="ARC6"
/ncbi_synonym="resulte in Glu to Asp substitution"
/replace="c"

ORIGIN

Alignment Scores:
Pred. No.: 1,31e-248 Length: 2438
Score: 4052.00 Matches: 799
Percent Similarity: 99.9% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.7% Indels: 1
DB: 15 Gaps: 0

US-10-600-070B-2 (1-801) x AY221469 (1-2438)

Cy 1 MetGUAleuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
|||||
Db 2 ATGGAGCTCTGAGTCAAGTCAGCTCGGCAATTGGCTCTCTCCCATTCCAATTATGCCGATTACCA 61
21 ProAlaThrThrIyLeuArgArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
|||||
Db 62 CCGGGAGCAGCAAAAGCTCCGAGGTGACCAACAACCTCATCAACAATATGCTCGCGCAGC 121
41 LysTTPAlaAspArgLeuLeuSerAspPheAsnPheMetThrSerAspSerSerSerSer 60
|||||
Db 122 AAATGGGGCGGACCGCTTCTCTCCGACATTCACCTCCGATTCCTCTCTCCCTCC 181
61 PheAlaThrAlaThrThrThrAlaThrIleuValSerLeuProProSerIleAspArgPro 80
|||||
Db 182 TTCGCACCGCCACCAACACCGCCACTCTCGTCTCTTCGCACCAACATATGATGTGCC 241
81 GluArgHisValProIleProIleAspPheArgIleValLeuGlyValGlnThrHisPhe 100
|||||
Db 242 GAAGCGCACGTCCTCCATCCCATGATTCTTCAACAGATATTAAGAGCTCAAAACATATTC 301
101 LeuThrAspGlyIleArgArgAlaPheGlnValaArgValSerIyGProProGlnIlePheIy 120
|||||
Db 302 TTAAACCGATGGAATCAGAAAGACATTCGAAAGCTAGAGTTTCCAAACCGCCGCAATTCGAT 361

```

QY	121	PheSerAapAapAalAleuLlIseArArgarGinlIleuGlnAlaAlaCybGluThrLeu	140
Db	362	TTcAGCGcAGcAGcCTTTATcACCGCGAGcAGATcCTTcCAAGcCTTGCGAAAcCTCG	421
QY	141	SerAapProAaGSeSerArgAgluThyrAenGluGlyLeuLeuAapAapGluGluAlaThr	160
Db	422	TCtAAATcCTCGGCTcTAAGAGAGATcCAATGAAGGCTcCTTGATGATcAGAGAcCTAc	481
QY	161	VallIethrAapValProTrpAapLysValProGlyAlaLeuCybValLeuGlnGluGly	180
Db	482	GTcATcAcTGAAGnGTcCTTGGAATAAAGTTcCTGGGGcCTcCTGcGTATTGcCAAGAGc	541
QY	181	GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys	200
Db	542	GGTGAGAcTGAAGATGTTcCTCGGGcTGcGTGAGcGCTcCTTAAAGAGAGcGTTGcCTTAAG	601
QY	201	SerPheLysGlnAapValLeuValMetAlaLeuAlaPheLeuAapValSerAaGAP	220
Db	602	TCGTTTAAAGCAAGATGcGTGTtTAAGTAAAGGcGCTGcCTTCTcCAATGCTcCGAGGAT	661
QY	221	AlaMetAlaLeuAapProProAapPheIleThrGlyTyrgluPheValGluGluAlaLeu	240
Db	662	GCtAAGcATtGGAATcCAcCTGATTTATtAcGTGTtAAGATtTGtGAGAAcCTTG	721
QY	241	LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAapLeuArgAlaGlnIleAap	260
Db	722	AAAGcTTTTTAAcAGAGAGAAAGcCAAGTAcCTTGcACCGcGATTTAcGTGCACAAATtGAT	781
QY	261	GluThrLeuGluGluIleThrProArgTyrgluLeuGluLeuLeuGlyLeuProLeuGly	280
Db	782	GAGAcCTTGGAAGAGATcCTCGGcGTATGCTTGAGAcCTAcTTGGcCTTAcCGcCTTGcGT	841
QY	281	AAPaAPTyrgluAlaLysArgLeuAenGlyLeuSerGlyValArgAanIleLeuTrpSer	300
Db	842	GATGATtTAAGcCTGCcCAAAAGcCTAAATGcTTTAAAGcGcGTGcGAAATtTTTGcGTCT	901
QY	301	ValGlyGlyGlyAlaLysAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAen	320
Db	902	GTTGAGAGAGcGTGAGAcATcCAcCTcCTTGcGGGcGTTTAcCCGTcAGAAcGTTTAAAT	961
QY	321	GluAlaPheLeuAaArgMetThrAlaIleGluGlnValAapLeuPheValAlaThrProSer	340
Db	962	GAGGcGcTTTTTAAcGAATGACAcCTGcGTGAGcAGcGTGATcTTTGAGcCTAcCCAAAGc	1020
QY	341	AanIleProAlaGlySerPheGluValTyrgluValAlaLeuAlaLeuValAlaGlnAla	360
Db	1022	AAATATcCCAGcAGAcATcCTTGAAGTTcAGAAcGTTGAcCTTGcCTTGcGcCTAAAGcT	1080
QY	361	PheIleGlyLysLysProHlaleuLeuGlnAapAlaAapLysGlnPheGlnGlnLeuGln	380
Db	1082	TTTATtTGATAAAGcCAcAcCTTTTAAcAGATcGTcGATAAAGcCAATtCCAGcAAcCTTcAG	1141
QY	381	GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrgluPheThrArgAanAenTrp	400
Db	1142	CAGGcCTAAAGcTAATGcCTATGAGATtCTGcCATtTGATATcAAcCGAAATtAAATtGG	1200
QY	401	GluIleAapPheGlyLeuGlnArgGlyLeuCybAlaLeuLeuIleGlyLysValAapGlu	420
Db	1202	GAAGATGAcCTcGcGTtTAAGAAAGGGAcCTcGTGcAcCTGcTTATtAGcCAAAcGTTGAAGAA	1260
QY	421	CybaArgMetTrpLeuGlyLeuAapSerGluAapSerGlnTyrgluAapProAlaIleVal	440
Db	1262	TGCcGATcGTGGcGTGGAcCAcGTGAcGTAGcGATtCCAATATtAGGAATcCAAGcCTATtGTG	1321
QY	441	GluPheValLeuGluLysAenSerAanAlaGluAapAanAapAapLeuProGlyLeuCybLys	460
Db	1322	GAGTTGTtTTGGAGATtCCAATcGTATGACAAATGAATGATcCTCCcTGAcCTATGcAAA	1381
QY	461	LeuLeuGluThrTrpLeuAlaGlyValAlaPheProArgPheArgAapThrLysAapLys	480
Db	1382	TTGTtGGAAAcCTcGcGTtGGcAGGcGcGTtGCTcTTCTTAAGcCTcAGAGcACcCAAAAGATtAA	1440
QY	481	LysPheLysLeuGluLysTyrgluTyrgluAapProMetValLeuSerTyrgluGluArgVal	500

[illegible]

VERSION	KEYWORDS
AY150490.1	GI:23297711
SOURCE	ORGANISM
Arabidopsis thaliana (thale cress)	Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	AUTHORS
1 (bases 1 to 2436)	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Quach, H.L., Tang, C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carrinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Timp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
2 (bases 1 to 2436)	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carrinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Timp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	JOURNAL
Arabidopsis Open Reading Frame (ORF) Clones	Unpublished
REFERENCE	AUTHORS
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA') (Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carrinci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.
COMMENT	JOURNAL
The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Timp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.	
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.	
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.	
FEATURES	source
location/qualifiers	1..2436
/organism="Arabidopsis thaliana"	/mol_type="mRNA"
/db_xref="taxon:3702"	/chromosome="5"
/clone="U19395"	/ecotype="Columbia"
/note="This clone is in pUNI 51."	1..2436
/gene="At5g42480"	1..2406
/gene="At5g42480"	/codon_start=1
/evidence=experimental	/product="unknown protein"
/protein_id="AA12907.1"	/db_xref="GI:23297712"
/translation="MEALSHVIGCLSPQCRULPATYKLRASNTSTTCSAKMADRLSDPSTDSSTSSSTRAITATITLTVSPSPSIDRPRHPYIPIDFQVJGAGQHFITDGIARRAEPARSKPQGFSDDALISRQILQACETLSPRSRRREYNGLIDDEKNTVTIDPMDKPVGALCVLQEGGTEIIVLRVGBALKEKLPSFKQDVVLAVALFADVSRLDAAALPDPETITGEYVEEALKLQEGGSLAPDLRAQIDETLEITRYYVELGLIPGDDVAARKINGLISGVRLNWSVGGGASALVGLTRKFPNBEATLPMYAEQVDLFTATSPNIPASFEVYVALALVAQATIGKKPHLLQDADNQFQQLQQAQKMMAMEIPAM	

LVDTRNNWEIDFGLERGLCALLIGKVDCEKMWLGLDSEDSQYRNPALVFEVLNSNRD
 DNDLPLGLCLLETWLAGVFPFRFDKDKKFLGDYDDPWLSTYLERFEVQGPPL
 AAAAAARIGAEHVKASAMAKOKFEPSPADRNASAPKOVQETVSVDPGNVRD
 GERGVFAEVRPSENENETNDVAIRAGVSSVDETTVEASVDMKEASVKILIAAGV
 AIGLISLPQKVPFLKSSSPORDMVSMSDVATIGSVADNSELPRDADATANI
 VSRKQIKSLAFEPDRHREMLPEVLDGRMLKITDRALAEIQAQGLYDTYLLKLSIDS
 VTVSADGTRALVEATLEESACLSDLVHPENNADIVRTYTRVEFPWSKGMKTGESV
 LAS"

3'UTR
 /gene="Acs542480"
 misc_difference 2433.2434
 /gene="Acs542480"
 /note="compared to genomic sequence"
 /replace="ctg"

ORIGIN

Alignment Scores:

Pred. No.:	1,52e-248	Length:	2436
Score:	4051.00	Matches:	799
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	2
Query Match:	99.7%	Indels:	0
DB:	15	Gaps:	0

US-10-600-070b-2 (1-801) x AY150490 (1-2436)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 Db 1 ATGGAGCTTGATGATCGTCGCGATTGCTCTCCCATTCATTTATTCGCAATTCACA 60
 QY 21 ProAlaThrThrIysLeuArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
 Db 61 CCGGCGACGACAAAGCTCCGACGTAAGCACAACTCTACAACTATCTCTCGCGACG 120
 QY 41 LysETPAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 Db 121 AAATGGCCGACCGCTCTCTCTCGACTTCATTTTCACTCCGATTCCTCTCTCTCC 180
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 Db 181 TTGGCCACCGCACACACCGCACCTCTGCTCTCCGCCACCATTTATTCATTCGCTCC 240
 QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100
 Db 241 GAACGCCACGTCCTCCATCCCATTTGATTTCTACACAGGATTTAGAGGCTCAAAACATTC 300
 QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIysProProGlnPheGly 120
 Db 301 TTAAACGATGGAATCAGAAAGCAATTCGAAGCTAGGGGTTTCGAAACCGCGCAATTCGGT 360
 QY 121 PheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
 Db 361 TTACGACACGACGCTTTAATCAGCCGAGACAGATTTTCAAGCTGCTTCGAAACCTCTG 420
 QY 141 SerAspProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspAspGlnGlnAlaThr 160
 Db 421 TCTAATCTCGCTAGAAAGAGTACATGAAGGCTCTTCTGATGATGAAGAAGCTACA 480
 QY 161 ValIleThrAspValProTTPAspLysValProGlyAlaLeuCyValLeuGlnGlnGly 180
 Db 481 GTCATCATCATGATCTCTTGCGATTAAGGTTCTGCTGCTCTGATTAAGCAAGAGGT 540
 QY 181 GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys 200
 Db 541 GGTGAGCTAGATAGTCTTCGGGTGTGAGGCTCTGCTTAAGGAGAGGTTGCTTAAG 600
 QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 Db 601 TCGTTTAAGCAAGATGTGTTTATGATTAGCGCTTCGTTTCTCGATGTCTGAGGAGAT 660
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
 Db 661 GCTATGGCATTTGATCCACTGATTTTATTAACGTGTTATGATGTTTGTGGAAGGCTTTG 720

QY 241 LysLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
 Db 721 AAGCTTTTACAGGAGGAG 780
 QY 261 GluThrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 Db 781 GAGACTTTGGAAG 840
 QY 281 AspAspTyrAlaAlaIleAspArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSer 300
 Db 841 GATGATTAAGCTGCGAAG 900
 QY 301 ValGlyGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGlnLysPheMetAsn 320
 Db 901 GTTGAG 960
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
 Db 961 GAGGCTTTTTCGAAATGACAGCTGCTGACAGAGTTGATCTTTTGTAGCTACCCCAACG 1020
 QY 341 AsnIleProAlaGluSerPheGlnValTyrGluValAlaAlaLeuAlaLeuValAlaGlnAla 360
 Db 1021 AATATTCCAGCAGAGCTCATTTGAGATTTCGAAATTTCGACTTCTTGTGGCTCAAGCT 1080
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnGlnGln 380
 Db 1081 TTTATTGTATGAAGAGCACACCTTTTACAGAGATGCTGATACGAAATTCAGCAACTTCAG 1140
 QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTyr 400
 Db 1141 CAGGCTTGAAGTATGCTATGAGATCTCGCATGTTATGATACAGGAAATTAATTTG 1200
 QY 401 GluIleAspPheGlyLeuGlnGlnArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
 Db 1201 GAGATGACTTCGCTTTAGAAAGGAGACTCTGTGCACTGCTTTATAGGCAAGTTGATGA 1260
 QY 421 CysArgMetTyrLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
 Db 1261 TGGCGATGTGGTGGCTTAAACATGAGATTCACAAATTAAGAAATTCAGCTAATTTG 1320
 QY 441 GluPheValLeuGlnLysSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
 Db 1321 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAAGATGATCTCTGACATATTCAA 1380
 QY 461 LeuLeuGluThrTyrLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
 Db 1381 TGTGTGAAACCTGCTTGGCAGGGGTTGCTTCTTCTAGGTTGAGACACCAAGATTA 1440
 QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGlnArgVal 500
 Db 1441 AATTTTAACTGGGGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
 Db 1501 GAGGTATGTCAGGGTCTCTCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
 Db 1561 CATGTAAAGCTAGTCTATGACGAGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
 Db 1621 AGAAATTCGGCTGAACCAAGAGATGCAAGAGACAGTGTATTAGTATGATCCTGTTGGT 1680
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 Db 1681 AACCAATGTAGGCGCTGATGATGAGCTGCTGCTTATTAATGCAAGACTTAACACCTCT 1740
 QY 581 GluAsnPheGluThrAspAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 Db 1741 GAATACTTGAACCTATATGATTTATGCAATTCAGCTGGGCTCTCAAGAGATGACGTTGAT 1800

Qy 601 GluThrThrValGluMetSerValAlaAspMetLeuValGluAlaSerValIleLeu 620
 Db 1801 GAAACATCTCTTGAATATCGCTTCTGATATGTTAAAGAGGCAATGTGAATCTTA 1860
 Qy 621 AlAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleValPheLeuLeuSer 640
 Db 1861 GCTGCTGGTGTGGCAATGCACTGATTTCACTGTTCCAGCCAGAAATTTTCTTAAAGC 1920
 Qy 641 SerSerSerPheGlnValGlyValAspMetValSerSerMetGluSerAspValAlaThrIle 660
 Db 1921 AGCTCATCTTTTCAACCAAGAAATATGTTCTTCTTGAATCTGATGTCGTACACATA 1980
 Qy 661 GlySerValAlaAlaAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 Db 1981 GGGTTCAGTCAAGCTGACGATTCAGAAACCTTCCAGATGATGCTAGACCTGCAGAG 2040
 Qy 681 AsnIleValSerIleValTrpGlnIleValIleValSerLeuAlaPheGlyProAspHisArgIle 700
 Db 2041 AATATAGTATCCAAAGTGCAGAAAGATTAAAGTCTCTGCTTTTGGGCTGATCACCGCAT 2100
 Qy 701 GluMetLeuProGluValLeuAspGlyArgMetLeuValIleTrpThrAspAlaAla 720
 Db 2101 GAAATGTATACCAAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGAGCAGCT 2160
 Qy 721 GluThrAlaGlnLeuGlyLeuValIleValAspValThrLeuLeuValSerValAspSer 740
 Db 2161 GAAACTCCGACCTGGGCTTGTATGATTAATACCTGTTGAACCTATCTGTTGACAGCT 2220
 Qy 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGlnIleValSerAla 760
 Db 2221 GTGACAGTCTCAGACAGATGAAACCCGCTCTGCTGAGAACCACTCTGAGAGAGTCTGCT 2280
 Qy 761 CysLeuSerAspLeuValHisProGluAspAsnAlaThrAspValAlaGlyThrIleThr 780
 Db 2281 TGCTATCTGATTTGGTTCATCCAGAAACATGCTTCTATGTCAGAACCTTACACACA 2340
 Qy 781 ArgTyGluValPheTrpSerIleValSerGlyTrpIleValThrGluGlySerValLeuAla 800
 Db 2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAATCTGTAAGGCTCTGTTCTTGCA 2400
 Qy 801 Ser 801
 Db 2401 TCA 2403
 RESULT 3
 LOCUS AY091075 2637 bp mRNA linear PLN 18-SEP-2002
 DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
 ACCESSION AY091075
 VERSION AY091075.1 GI:20259550
 KEYWORDS FLU CDNA.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana (chale crees)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 2637)
 Yamaoka, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 Unpublished
 2 (bases 1 to 2637)
 Yamaoka, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission
 Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP cDNAs (RAFL cDNA; 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamaoka, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamaoka, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PI.
 Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
 Location/Qualifiers
 1. 2637
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL09-76-G11 (R19395)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector (pUC-1) as a BamHI/XhoI insert."
 1. 2637
 /gene="At5g42480"
 1. 113
 /gene="At5g42480"
 114. 2519
 /gene="At5g42480"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AA13895.1"
 /db_xref="GI:20259551"
 /translation="MRALSHVIGLSPPOLCRLPATTKLRBSHNTTTCASKNADRLSDPNFTSDSSSPATRTTATTIVSSPPSIDRBRHPRTIDRYQUGATHTPLTGIRRAFEARVSKRPQGFSDDLISRRQLQAACSTLSPRSRRYBGLDDEKATVITDVPMDKVPALCVLOSGGTEIYLRVGEALIKRLPSPSQDVLVVALAFLDVS RDMAALDPDPITGVEFEVALGLDSEGSALPDLRAIDITLEITFRVYELLG LPIGDYTAABRLNGLSGVGSLVGLTEKFNBAFLRWTAAEYDL FVATPNTIABRSFEYEVALLVAQAFLKKPHLLDDAKQPOQLQAATMAEIPAM LVTRNNWEIDFGLERGLCALIGKYDECKMGLDSEDSQIRNPATIVERVLNSND DUNDLPGLCLLETWLAGVFPFRFDTKKFKGLGYYDPMVLSYLEREVVQGSPL AAAAAMARACAEHVAKASAAQLOKVPSPSYTDSNAPKOVQETFSVDPGVNNGVD GEGGVIAEVRSENFENDNVAIRAGVSSVDETVEMSVADMKEASVKTLIAGV ATGLISLPSOKYFLKSSSFNDYATGVSMSVSDYATGVSRAADSEALPRMDARTAEI VSKWQKIKLAREPDIRIEMLEPVLDGRMLKLTPTDAAETAOGLVYDITLGLSDS VVVSAGCTALVEATLESACUSDVLPENNATDVRTYTRIEVFSKSGMKITBGSV LNS"
 2637
 /gene="At5g42480"
 ORIGIN
 3'UTR
 Alignment Scores:
 Pred. No.: 1,68e-248 Length: 2637
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2

Query Match: 99.7% Indels: 0
DB: 15 Gaps: 0
US-10-600-070b-2 (1-801) x AY091075 (1-2637)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCybArgLeuPro 20
DB 114 ATGAAGCTCTGATGACGTGGGCAATGATCTCTCCCATTCACAAATATATCCCATTAACA 173

QY 21 ProAlaThrThrIlyLeuAArgArgSerHisAsnThrSerThrTrilleCysSerAlaSer 40
DB 174 CCGGCGACGCAAAAGCTCCGACGTAAGCACAAACCTCTACAACTATCTCTCCGCGACG 233

QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPhenThrSerAspSerSerSerSer 60
DB 234 AAATGGGCGCAACGCTCTCTCTCCGACTTCAATTTCCATTCCTCCGATCTCTCTCTCC 233

QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
DB 294 TTGCGCACCGCCACACACCGCACTCTGCTCTCCGCAACCATCTATTTGATCGTCCC 353

QY 81 GluArgHisValProIleProIleAspPheTyrGlnIleGluValaGlnThrHisPhe 100
DB 354 GAAAGCGACGTCCCGCATCCCATTTGATTTCTACCGAGTATTAAGAGCTCAAAACATTTCC 413

QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
DB 414 TTAAACGATGGAATCAGAAAGCATTCGAAAGCTTGGGTTTCGAAACCGCGCAATTCGT 473

QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleGluAlaAlaCysGluThrLeu 140
DB 474 TTCAGGACGACGCTTATTCAGCCGAGACAGATTTCTTCAAGCTGCTTCCGAAACTCTG 533

QY 141 SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluAlaThr 160
DB 534 TCTAATCTCGGTCTAGAAAGAGTACAAAGAAAGTCTTCTGATGATGAAGAAAGCTACA 593

QY 161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuCysValLeuGlnGluGly 180
DB 594 GTCACTCATGATGTTCTTGAGATTAAGTCTCTGCTCTCTCTGATATTCAGAAAGGT 653

QY 181 GlyIleThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys 200
DB 654 GGTGAGACTGAGATGATTTCTTCCGGTGTGGAGGCTCTGCTTAAAGAGAGTTCCTTAAG 713

QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
DB 714 TCGTTTAAGCAAGATGTGTTTATGATAGCGCTTCCGTTTCGATGTCTCGAGGGAT 773

QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
DB 774 GCTATGCAATTTGATCCACTGATTTTATTAAGTGTATGATGATTTGTTGAGAAAGCTTTC 833

QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
DB 834 AACCTTTTACAGAGGAGAGCAAGTAAAGCTTTCACCGGATTTAAGTGAACAAATTGAT 893

QY 261 GluThrLeuGluGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGly 280
DB 894 GAGACTTTGAGAGATCACTCCGCTTATGTCTTGAAGCTACTTGGCTTAACCCCTTGGT 953

QY 281 AsnAspTyrAlaIleLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
DB 954 GATGATTAACCTGCGAAAGACTTAATAGTTTAAAGCGGTGCGGAAATTTTGTGTCT 1013

QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
DB 1014 GTTGGAGAGGTGAGATCACTCTTGTGGGGTTTGAACCGTGAAGATTATGAT 1073

QY 321 GluAlaPheLeuArgMetThrAlaIleGluGlnValAspLeuPheValAlaThrProSer 340
DB 1074 GAGCGTTTATACAAATGACAGCTGTGAGAGGTTGATCTTTTGTATGCTACCCCAAGC 1133

QY 341 AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValaGlnAla 360
DB 1134 AATATTCCAGACAGAGTCATTTGAAGTTTACGAAGTTCACATTGCTCTTGTGGCTCAAGCT 1193

QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
DB 1194 TTTATTTGTAAGAGCACACCTTTTACAGAGTCTGATTAAGCAATTTCCAGCACTTCAG 1253

QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrp 400
DB 1254 CAGCTTAAGTAAATGCTATGAGATTCCTGCGATGTTTATGATACAGGAATATTTGG 1313

QY 401 LysLeuAspPheGlyLeuGluArgGlyLeuCysAlaLeuLeuIleGlyLysValaArgGlu 420
DB 1314 GAGATAGACTTCGGTCTTAAGAAAGGAGACTGTGCACTGCTTATAGGCAAAAGTTGATGA 1373

QY 421 CysArgMetTrpLeuGlyLeuAspSerGlyLysAspGlnTyrArgAsnProAlaIleVal 440
DB 1374 TGCCTATGTGTGGGCTTGAACAGTGAAGATTCAAATATAGGAATCCAGCTATTGTG 1433

QY 441 GluPheValLeuGluLysAsnSerAsnArgAspAsnAspLeuProGlyLeuCybLys 460
DB 1434 GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGAGCTATGCCAA 1493

QY 461 LeuLeuGluThrTrpLeuAlaGlyValaIlePheProArgPheArgAspThrLysAspLys 480
DB 1494 TTGTTGAAAACCTGTTGGCAGGGGTTGCTTCTTCTTCAAGTTTCAGAGACACAAAGATTA 1553

QY 481 LysPheLysLeuGluLysAspTyrTyrAspAspProMetValLeuSerTyrLeuGluArgVal 500
DB 1554 AAATTTAACTCGGGACCTATGATGATCCATATGTTTGAAGTACTTGGAAAGAGTG 1613

QY 501 GluValValaGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGln 520
DB 1614 GAGGTATCTCAGGGTCTCTCTTATGCTGCTGCTGCGACATGCAAGATTTGAGCGGAG 1673

QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
DB 1674 CATGTAAAGCTAGTCTATGACGACCTGCAAGAAAGTTTTCCTCCGCTATACAGAT 1733

QY 541 ArgAsnSerAlaGluProLysAspValaGlnGluThrValPheSerValAspProValaGly 560
DB 1734 AGAACTCGGCTGAACCAAGATGTGCAAGACAGTGTTAAGTATGATCTGTTGGT 1793

QY 561 AsnAsnValaGlyArgAspGlyGluProGlyValaPheIleAlaGluAlaValaArgProSer 580
DB 1794 AACATGTAGGCGGTATGATGATGCTGCTGCTTATATTCAGAACTGTAAAGCCCTCT 1853

QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValaSerGluSerSerValaAsp 600
DB 1854 GAATACTTGAACCTAATGATTTATGCAATTCGAGCTGGGGTCTCAAGAGTACGTTGAT 1913

QY 601 GluThrThrValaGluMetSerValaAlaAspMetLeuLysGluLysSerValaLysIleLeu 620
DB 1914 GAACCTACCTGTGAATGTCCGTTGATATGTTAAAGAGGCAAGTGAAGATCTTA 1973

QY 621 AlaAlaGlyValaAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSer 640
DB 1974 GCTGCTGGGTGGCAATTTGACTGATTTCACTTACGTCAGCAAGATTTTCTTAAAGC 2033

QY 641 SerSerSerPheGlnArgLysAspMetValaSerSerMetGluSerAspValaAlaThrIle 660
DB 2034 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATGGAATCTGATGTCGCTACCATTA 2093

QY 661 GlySerValaArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGln 680
DB 2094 GGGTCAGTCAGACCTGACATTCAGTCAAGACACTTCCAGATGATCTAGAGCTGACAG 2153

QY 681 AsnIleValaSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgGly 700
DB 2154 AATATATGATCCAAAGTGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2213

QY 701 GluMetLeuProGluValaLeuAspGlyArgMetLeuLysIleTrpThrAspArgAlaAla 720

Dd	2214	GAAATCTTACGAGGTTTGATGGCCGAATGTGAAGTTTGATGCACAGCAGCT	2273
Oy	721	GIUThrAlIGlnLeuGlyLeuValTYrAspTYrThrIeuLeuLysSerValAspSer	740
Dd	2274	GAAACTGCCACAGCTGGGTGGTTTATGATTATACTGTTGAAACTATCTGTTGACAGT	2333
Oy	741	VAlTHrValSerLAspRgiYThrArgAlaLeuValGluValaThrLeuGluGluSerAla	760
Dd	2334	GGACAGCTTCAGCAATGGACCCTCGCTGGTGAMCAACTCTGGAGGAGTCTGCT	2393
Oy	761	CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgTrpTYrThr	780
Dd	2394	TGTCrATCTGATTGGTTGTTTCATCCAGAAAACAAGCTAATGTCAGAACCTACACA	2453
Oy	781	ArGYrGluValPheITrpSerLYrsERGlyTRPLysILEThrgluGlySerValLeuAla	800
Dd	2454	ACATACGAAGTTTCTGCTCAAGTCAAGCGGTGGMAATCATGTAAGGCTCTGTTCTTGCA	2513
Oy	801	Ser 801	
Dd	2514	TCA 2516	
<hr/>			
RESULT 4			
LOCUS	AY221468	3668 bp DNA linear	PLN 05-AUG-2003
DEFINITION	Arabidopsis thaliana division protein (ARC6) gene, complete cds;		
ACCESSION	AY221468		
VERSION	AY221468.1	GI:33436338	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.		
REFERENCE	1 (bases 1 to 3668) Vitsha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Epp,H. and Osteryoung,K.W. ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2 Plant Cell 15 (8), 1918-1933 (2003)		
AUTHORS	2 (bases 1 to 3668) Vitsha,S., Koksharova,O., van Epp,H., Froehlich,J.E. and Osteryoung,K.W. Direct Submission Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA Location/Qualifiers 1..3668 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /ecotype="Massilewskija" 478..3515 /gene="ARC6" /note="synonym: AtSg42480" join(<478..984,1067..1291,1379..1639,1723..2710, 3000..3133,3225..>3515) /gene="ARC6" /product="division protein" join(478..984,1067..1291,1379..1639,1723..2710,3000.. 3225..3515) /gene="ARC6" /note="inner envelope membrane-localized; similar to cyanobacterial cell division protein slt10169 of Synechocystis; contains J-domain" /codon_start=1 /product="division protein" /protein_id="AAO18645.1" /db_xref="GI:33436339" /translation="MEALSHVGIGSLFPCRLRPATTKLRSHNTSTTICSAKMD		

[illegible]

QY 213 aPhelLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyr 233
 DB 1198 GTTTCGATGCTCTCCAGGAGATGCTATGCGATTCGACCTGATTTTATTTACTGCTTA 1257
 QY 233 rGluPheValGluGluAlaLeuValLeuGln----- 244
 DB 1258 TGAGTTGTTGAGGAGCTTTGAAGCTTTTACAGGTAGTTGACTTCTTGGAATTTTG 1317
 QY 244 ----- 244
 DB 1318 ACCAGCGTTGGCTTTATAGAACTTTCTGATTTGATTACTTTGTTATTTAGCTTGCTGA 1377
 QY 245 -GluGluGluValSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGlu 264
 DB 1378 GGAAGAGAGAGCAAGTACCTTGACACCGGATTTACGTCGACAAATTGATGAGCTTTGGA 1437
 QY 264 uGluIleThrProArgTyrValLeuGluLeuGluLeuGluLeuProLeuGluValAspPheTyrAl 284
 DB 1438 AGAGATCACTCCGCGTTATGCTTGAGACTTGGCTTACCGCTTGATGATGATTTACGC 1497
 QY 284 aAlaValArgLeuAsnGluValLeuSerGluValArgAsnIleLeuThrSerValGluValGlu 304
 DB 1498 TGGGAAAAGACTAAATGCTTTAAGCGGTGTGCGGAATATTTTGTGTGTGTGAGAGAG 1557
 QY 304 yGluValSerAlaLeuValGluValLeuThrArgGluValPheMetAsnGluAlaPheLe 324
 DB 1558 TGGAGCATCAGCTCTTGTGGGGGTTCACCCGAGAGATTTATGATGATGAGGCTTTT 1617
 QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
 DB 1618 ACCAATGACAGCTGCTGAGCAGGT-ATACAGTTTATGATACCTTTTAAATTTCTTTAGC 1676
 QY 333 -----AspLeuPheVal 336
 DB 1677 ATGATATACTTAGGTTTCTCATTTTAATGATATGTTGTGTGTAGGTATCTTTTGT 1736
 QY 336 lAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLe 356
 DB 1737 AGCTTACCCCAAGCAATTTCCAGCAGAGTCACTTTGAAGTTTACGAAATTCATCTGCTCT 1796
 QY 356 uValAlaGlnAlaPheIleGluValValPheProAlaLeuGluAlaAspValAspValGln 376
 DB 1797 TGTGGCTCAAGCTTTTATGTTAGTAAAGCCACACTTTTACAGGATCTGTATTAAGCAAT 1856
 QY 376 eGlnGlnLeuGlnGlnAlaValMetAlaMetGluIleProAlaMetLeuTyrAspThr 396
 DB 1857 CCAGCAACTTCAGAGGCTAAGTATGCTATGAGATTCCTGCGATGTTGATGATAC 1916
 QY 396 rArgAsnAsnTropGluIleAspPheGluValLeuGluValGluValLeuGluValLeuGlu 416
 DB 1917 ACCGAATATTTGGAGATGAGTCTCGATCTAGAAAGGGAGCTCTGTGCACTGCTTATAGG 1976
 QY 416 yLysValAspGluCysArgMetTyrLeuGluValAspSerGluValAspSerGlnTyrArgAs 436
 DB 1977 CAAGTGAATGAGATGCCGTATGTGTTGGGCTTATGACAGTGAAGATTCACAATATAGAA 2036
 QY 436 nProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAsnAspAspLeuPhe 456
 DB 2037 TCCAGCTATTTGTGAGTTTGTGTTGGAAATTCAAATCGATACATATGATGATCTCC 2096
 QY 456 oGluLeuValLeuLeuGluGluThrTyrPheAlaGluValValPheProArgPheArgAs 476
 DB 2097 TGGACTATGCAAAATTTGGAACCTGTGTGGACGGGGGTGTCTTTCTTACGTTCAAGAA 2156
 QY 476 pThrLysAspLysLysPheLysLeuGluValAspTyrTyrAspAspProMetValLeuSerTyr 496
 DB 2157 CACCAAGATTAATAAATTTAACTCGGGACTACTATGATGATGATCTTATGTTTGAAGTTA 2216
 QY 496 rLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaThr 516
 DB 2217 CTGGAAAGAGTGAAGTGTAGTTCAAGGGGTCTCTTATGCTGCTGCAACTATAGGCAAG 2276

QY 516 gIleGluAlaGluHisValIleValSerAlaMetGlnAlaLeuGlnLysValPheProSe 536
 DB 2277 GATTGAGCGGAGCAATGTGAAGCTATGCTATGACGAGCACTGCAAGAAAGTTTCTCTTC 2336
 QY 536 rArgTyrThrArgAspAsnSerAlaGluProLysAspValGlnGluThrValPheSerVal 556
 DB 2337 CCGCTATTAACAGATGAAGAACTCGGCTAACCCAAAGATGAGCAAGACAGCTTTTAGTGT 2396
 QY 556 lAspProValGluValAsnValGluValArgAspGluGluProGluValPheIleAlaGluAl 576
 DB 2397 AGATCTGTTGGTGAACATGATGAGCGGTATGCTGAGCGCTGCTTATTTATTCAGAGAC 2456
 QY 576 aValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGluValSerGlu 596
 DB 2457 TGTAAAGCCCTCGAAAACTTTGAACCTATGATTTATGCAATTCGAGCTGGGGCTTCAGA 2516
 QY 596 uSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuValGluValase 616
 DB 2517 GAGTACCGTTGAGAACTACGTGTAATATGTCCTGTGATATGTTAAAGAGGCGAC 2576
 QY 616 rValLysIleLeuAlaAlaGluValAlaIleGluLeuIleSerLeuPheSerGlnTyr 636
 DB 2577 TGTGAAGATCTTACGCTGCTGCTGTGCGCAATTCGACTGATTCATCTGTTACGCAAGAGTA 2636
 QY 636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
 DB 2637 TTTTCTTAAAGCAGCTCATCTTTTCAACGACAGATATGTTCTTCTATGGAATCTGA 2696
 QY 656 pValAlaThrIle----- 660
 DB 2697 TGTGCTACCAATAGTATGATTAATGATGCAATTTTCATATATCTGATTCCTCAAAAT 2756
 QY 660 ----- 660
 DB 2757 ATGCTGTTTGTGACCTAAGAACATAGTCCCATTAATACATGTCCTCAAAAGTTGTAC 2816
 QY 660 ----- 660
 DB 2817 CAGATTAACAAGTGTGCTGATTAATTTCACTAATATGCTGCTGAATTTTGTGATCA 2876
 QY 660 ----- 660
 DB 2877 ACTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAGAAATTAAGTAGATTA 2936
 QY 660 ----- 660
 DB 2937 GAGATTCCTTATGATGCTGCTTGTCCAACTTTTCTTCTGATTTTCTTTTCTTGATTT 2996
 QY 661 --GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
 DB 2997 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTAGAGACTGAG 3056
 QY 680 lAsnIleValSerLysTyrGlnLysIleLysSerLeuAlaPheGluProAsnHisArg 700
 DB 3057 AGAATATAGATTCAGATGCGACAGAAATTAAGTCTGCTTGTGGGCTGATCACCGCA 3116
 QY 700 lGluMetLeuProGlu----- 705
 DB 3117 TAGAAATGTTTACAGAGGAGGAGAAATTAATCTTCAATTCATCAATTTGCTGAAAACTG 3176
 QY 706 -----ValLeuAspGlu 709
 DB 3177 TTGACATGATTAATAGCTGCTGCTTGTGATTCGTGTTATTTATAGGTTTGGATGGG 3236
 QY 710 ArgMetLeuLysIleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGluValValTyr 729
 DB 3237 CGAATGCTGAAGATTTTGACTGACAGAGCACTGGAACCTGCGAGCTTGCTGGTTGTTAT 3296
 QY 730 AspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGluThrArg 749
 DB 3297 GATTATACACTGTGGAACCTATCTGTTGACAGGTGACAGTCTCAGCAGATGGAACCCGT 3356
 QY 750 AlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGlu 769

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3664)
Vitha,S., Froehlich,J.E., Kosharova,O., Pyke,K.A., Van Erp,H. and Oteyyoung,K.W.

TITLE

ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2

JOURNAL

Plant Cell 15 (8), 1918-1933 (2003)

PubMed

12897262

2 (bases 1 to 3664)

REFERENCE

Vitha,S., Kosharova,O., van Erp,H., Froehlich,J.E. and Oteyyoung,K.W.

TITLE

Direct Submission
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA

JOURNAL

Location/Qualifier

FEATURES

1..3664

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/ecotype="Ws-2"

/note="genotype: arc6 mutant"

1..3664

/gene="arc6"

/allele="arc6-1"

/join(478..984,1067..1291,1379..>1621)

/gene="arc6"

/allele="arc6-1"

/product="truncated division protein"

/join(478..984,1067..1291,1379..1621)

/gene="arc6"

/allele="arc6-1"

/note="truncation caused by premature stop codon"

/codon_start=1

/product="truncated division protein"

/protein_id="AA018644.1"

/db_xref="GI:33436275"

/translation="MRALSHVIGLSPFOLCRLPATTKLRSHNTSTTCSAKMADRLISDFNFTSDSSSPATRTTATVSLSPSIDRPRRHPIRIDYQVGAOTHTLT

DGIRAFARVSRPQFGSDALISRLQAACETLSNPSRRRNEGLDDEBAT

VITRVPMDKRVGALCVLOEGEVEIVLRVEALIKERLPKSFQDVLVWALFLVVS

RDMAALDPDFITGEVEERALKILOEGASLAPDLRAQIDTLEITPTRVYLELIG

LPFGDDVYAAKRLNGLSGVNRIIWSVGGGASALVGVITREKFNNEAFL"

ORIGIN

Alignment Scores:

Pred. No.: 7,72e-234

Score: 3823.50

Percent Similarity: 78.7%

Best Local Similarity: 94.1%

Query Match: 15

DB: 15

US-10-600-070b-2 (1-801) X AY221467 (1-3664)

QY 1 Met(G)AAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
DB 478 ATGGAAGCTCTGAGTCAGTCGCGCATTCCTCTCCCATTCATTAATGCGCATTAACA 537
QY 21 ProAlaThrThrIleValLeuArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
DB 538 CCGGCGACGACAAAGCTCCGACGTAACCAACACCTTCACTAATCTGCTCCGACG 597
QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
DB 598 AAATGGGCGACGCTCTCTCTCGCACTTCAATTCACCTCCGATTCCTCCTCCTCC 657
QY 61 PheAlaThrAlaThrThrAlaThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
DB 658 TTCGCCACCGCCACACACCGCACCTCTGCTCTGCGCACCATCATTAATGATCGGCC 717
QY 81 GluArgHisValProIleProIleAspPheThrGlnValLeuGlyAlaGlnThrHisPhe 100

DB 718 GAACGCGACAGTCCCAATCCCATTTCTTCACAGATTTAGGACCTCAACACATTTTC 777
QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProPheGlnPheGly 120
DB 778 TTACCGATGGATCGAAGAGACATTCGAAGCTTAGGGTTTCCAAACCGCGCATTCGGT 837
QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
DB 838 TTCAGCGACGACGCTTAATTCAGCCGAGACAGATTCCTCAAGCTCTCCGAACTCTG 897
QY 141 SerAsnProArgSerArgArgIleValIleValIleValIleValIleValIleValIle 160
DB 898 TCTTAATCCCGCTGTAAGAGAGTCAATGAAGGCTCTTGATGATGAAGAAGCTTACA 957
QY 161 ValIleThrAspValProThrAsp 168
DB 958 GTCATCACTGATGTTCTCTGGATGAAGTAATTCGATTTCCGAATTAATTAAGTTCTTC 1017
QY 169 -----LysValProGlyAla 173
DB 1018 GTTTAATTTCAATGATGATGAATGAAGAAGAACTTTATCTAGTGAAGGTTCTCGGGC 1077
QY 173 AlaLeuValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 193
DB 1078 TCTCTGTATTCGCAAGAGAGTGTGAGACTGAGATAGTCTTCGGGTTGGAGGCTCT 1137
QY 193 uLeuLysGluArgLeuProLysSerPheLysGlnAspValValIleValIleValIleVal 213
DB 1138 GCTTAAGAGAGGTTCTCTTAAGTCTTTTAAGCAAGATGTGTTTATGTTATGCGCTTGC 1197
QY 213 aPheLeuAspValSerArgAspAlaMetAlaLeuAspProPheAspPheIleThrGlyTyr 233
DB 1198 GTTTCGATGTCGTGAGGAGTGTATGCAATTCGATTCACCGATTTATTACTGCTTA 1257
QY 233 rGluPheValGluGluAlaLeuLysLeuLeuGln----- 244
DB 1258 TGAGTTGTTGAGGAAGCTTTGAAGCTTTACAGTAGTTGACTTCTTGTTAATTGG 1317
QY 244 ----- 244
DB 1318 ACGAGGTTGGCTTTAAGAACTTTCTGATTTGATCTTTGTTATTCAGTCTGTGTA 1377
QY 245 -GluGluGlyAlaAspSerIleValProAspLeuArgAlaGlnIleAspGlyThrLeuGln 264
DB 1378 GAGGAAGAGACAAAGTACCTTGACCGGATTTACGTGCAAAATGAGAGACTTTGG 1437
QY 264 uGluIleThrProArgTyrValLeuGlnLeuLeuGlnLeuLeuGlnLeuLeuGlnLeuLeu 284
DB 1438 AGAGATCACTCCGCGTTATGTCTTGAGACTTACGCTTGACGCTTGAGATGATTAACG 1497
QY 284 aAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGly 304
DB 1498 TCGGAAGAAGCTAAATGTTTAAGCGGTGTGCGGAATATTTGTGCTGTGTGGAGAGG 1557
QY 304 yGlyAlaSerAlaLeuValGlyIleLeuThrArgGlyLysPheMetAsnGlyAlaPheLe 324
DB 1558 TGGAGATCACTCTCTGTTGGGGGTTTGACCCGTGAAAGTTATGAATGAGCGCTTTT 1617
QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
DB 1618 ATGAATGACAGCTGTGAGAGAGGT-ATACGTTAGATACCTTTTAAATTTCTTTAGC 1676
QY 333 -----AspLeuPheVal 336
DB 1677 ATGATTAACCTTAGGTTCTCATTTTAAATGATGTGTGTGATGAGCTTTTGTG 1736
QY 336 lAlaThrProSerAsnIleProAlaGlnSerPheGlnValIleGluValAlaLeuAlaLe 356
DB 1737 AGCTACCCCAAGCAATATTCACAGAGTCAATTTGAAGTTTACGAAGTTGCACTTCTCT 1796
QY 356 uValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPhe 376

Db 1797 TGTGGCTCAAGCTTTTATTGTTGTAAGAACACACCTTTTACAGGATGCTGTAAGCAATT 1856
 Qy 376 eGInGInLeuGInGInLaIyVaIMetAlaMerGInIlleProAlaMetLeuTyraAPh 396
 Db 1857 CCACCACTTACAGCAGGCTAAAGTAAGCTATGAGAGATTCCTCGATGTTGTAAGTAC 1916
 Qy 396 rArGAsnAsnTrpGluIleAspPheGlyLeuGluAArgGlyLeuCySaIaleuLeuIleG 416
 Db 1917 ACCGAAATATTGGAGATAGACTTCGCTGTAAGAAAGGGAGACTCTGTCACTGCTTAAG 1976
 Qy 416 yLyVaIaAspGluCySaIArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyraArg 436
 Db 1977 CAAGTGTAGTAAGCCGTAATGCTGTTGGCTTACAGTGAAGATTCACAATATAGGAA 2036
 Qy 436 nProAlaIleValGluPheValLeuGluAsnSerAspAspAspAspAspAspAspAsp 456
 Db 2037 TCCAGCTATTGTGAGTTGTTGTTGAGAAATTCAAATCGTATACAAATGATGATCTCC 2096
 Qy 456 oGlyLeuCyLeuLeuLeuGluTrpTrpLeuAlaGlyValValPheProArgPheArg 476
 Db 2097 TGGACTATGCAAAATGTTGGAACCTGCTGGCAGGGGTTCTTCTTCTAGGTTCAAGAA 2156
 Qy 476 pThrLyAspLysLysPheLysLeuGlyAspTyrrTyraAspProMetValLeuSerTy 496
 Db 2157 CACCAAGATTAATAAATTTAACTCGGGACTACTATGATGATGATCTATGTTTGAAGTTA 2216
 Qy 496 rLeuGluAArgValGluValValGlnGlySerProLeuAlaAlaAlaAlaAlaAlaAla 516
 Db 2217 CTGGAAAGAGTGAAGTGAAGTTCAGGGTTCTCTTGTAGCTGCTCAGCTAAGGCAAG 2276
 Qy 516 gIleGlyAlaGluIleValLysValLysAlaSerAlaMetGluAlaLeuGlnLysValPhePro 536
 Db 2277 GATTGAGCCGAGATGTAAGCTAGTCTATGCAAGCACTGCAAGAAAGTTTTCCTTC 2336
 Qy 536 rArgTyrrThrAspArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerVa 556
 Db 2337 CCCTATACAGATGTAAGCTCGGCTGAACCAAGATGTGCAAGACAGAGTTTACTGT 2396
 Qy 556 LaAspProValGlyAspAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAl 576
 Db 2397 AGATCTGTTGTAAGCTATGAGGCGGTGATGTGAGCCGTGCTTCTTATGTGAGAGC 2456
 Qy 576 aValArgProSerGluAsnPheGluThrAsnAspTyrrAlaIleArgAlaGlyValSerG 596
 Db 2457 TGTAAAGCCTCTGCAAACTTTGAAACTATGATTAATGCAATTCAGCTGGGCTCAGA 2516
 Qy 596 uSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSe 616
 Db 2517 GAGTAGGCTGATGAACCTACTGTTGAATGTCCGTTGCTGATGTAAAGAGGCAAG 2576
 Qy 616 rValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTy 636
 Db 2577 TGTGAAGATCTTGAAGCTGCTGCTGTGGCAATTCGACTGATTCCTGTTCAAGCAAGATA 2636
 Qy 636 rPheLeuLysSerSerSerSerPheGluArgLysAspMetValSerSerMetGluSerAs 656
 Db 2637 TTTTCTTAAAGAGCTCACTTTTCAACGAGATATGTTTCTTCTAAGATCTGA 2696
 Qy 656 pValAlaThrIle----- 660
 Db 2697 TGTGCTACCATAGGTATGATTAATATGATGCAATTTTCATATATCTGATTCGCAAAAT 2756
 Qy 660 ----- 660
 Db 2757 ATGCTGTTTGTGAGCTAAGAACATAGTCCCATTAATACATGTCCAAAGTTGTAC 2816
 Qy 660 ----- 660
 Db 2817 CAAGATTACAAAGTTGCTGAGTAATTTCACTAATATATGCTGCTGAATTTTGTATCAA 2876
 Qy 660 ----- 660
 Db 2877 ACTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAAGATAAGTAGATTA 2936

Qy 660 ----- 660
 Db 2937 GAGATTCCTTAGTGTGCTGCTTGTTCACACTTTCTTCTTGAATTTTCTTTCGATT 2996
 Qy 661 --GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
 Db 2997 TAGGTCAGTCAGAGCTGAGATTCACAGAACCTTCCCAAAATGATGATGATGCTCAG 3056
 Qy 680 LuAsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgI 700
 Db 3057 AGAATATAGTATCCAAATGTCAGAGATTAAGTCTGCTGTTGGCTGATCACCGCA 3116
 Qy 700 leGluMetLeuProGlu----- 705
 Db 3117 TAGAATGTATCCAGAGCTGAGGAAATTAATCTTAATCAATCAATGCTGTAAGAACTG 3176
 Qy 706 -----ValLeuAspGly 709
 Db 3177 TTGACATGATTAATAGTCTGCTGCTTGTGATTCGTTATTTATAGTTTGGATGGG 3236
 Qy 710 ArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTy 729
 Db 3237 CGAATGCTGAAGATTGGAAGTGAAGAGCAGAGCTGAAGAACTCGCAGCTTGGGTTGAT 3296
 Qy 730 AspTyrrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArg 749
 Db 3297 GATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCGT 3356
 Qy 750 AlaLeuValGluAlaThrLeuGluGluSerAlaCyLeuSerAspLeuValHisProGlu 769
 Db 3357 GCTCTGTGTAAGCACTCGAGAGAGTCTGCTGCTATGATTTGGTTCACTCCAGAA 3416
 Qy 770 AsnAsnAlaThrAspValaArgThrTyrrThrArgTyrrGluValPheTrpSerLysSer 789
 Db 3417 AACAAATGCTACTGATGTCAGAACCTACACAAAGATATACAAATTTCTGTCCAAGTCA 3476
 Qy 790 GlyTrpLysIleThrGluGluLysSerValLeuAlaSer 801
 Db 3477 GGGTGAATAATCACTGAAGGCTGTTTTCATCA 3512

RESULT 7
 AP008208_011/c
 WPCOMMENT
 Sequence split into 360 fragments LOCUS AP008208 Accession AP008208
 Fragment Name Begin End
 AP008208_000 1 110000
 AP008208_001 100001 210000
 AP008208_002 200001 310000
 AP008208_003 300001 410000
 AP008208_004 400001 510000
 AP008208_005 500001 610000
 AP008208_006 600001 710000
 AP008208_007 700001 810000
 AP008208_008 800001 910000
 AP008208_009 900001 1010000
 AP008208_010 1000001 1110000
 AP008208_011 1100001 1210000
 AP008208_012 1200001 1310000
 AP008208_013 1300001 1410000
 AP008208_014 1400001 1510000
 AP008208_015 1500001 1610000
 AP008208_016 1600001 1710000
 AP008208_017 1700001 1810000
 AP008208_018 1800001 1910000
 AP008208_019 1900001 2010000
 AP008208_020 2000001 2110000
 AP008208_021 2100001 2210000
 AP008208_022 2200001 2310000
 AP008208_023 2300001 2410000
 AP008208_024 2400001 2510000
 AP008208_025 2500001 2610000
 AP008208_026 2600001 2710000

AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	10110000
AP008208_101	10100001	10210000
AP008208_102	10200001	10310000
AP008208_103	10300001	10410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000
AP008208_132	13200001	13310000
AP008208_133	13300001	13410000
AP008208_134	13400001	13510000
AP008208_135	13500001	13610000
AP008208_136	13600001	13710000
AP008208_137	13700001	13810000
AP008208_138	13800001	13910000
AP008208_139	13900001	14010000
AP008208_140	14000001	14110000
AP008208_141	14100001	14210000
AP008208_142	14200001	14310000
AP008208_143	14300001	14410000
AP008208_144	14400001	14510000
AP008208_145	14500001	14610000
AP008208_146	14600001	14710000
AP008208_147	14700001	14810000
AP008208_148	14800001	14910000
AP008208_149	14900001	15010000
AP008208_150	15000001	15110000
AP008208_151	15100001	15210000
AP008208_152	15200001	15310000
AP008208_153	15300001	15410000
AP008208_154	15400001	15510000
AP008208_155	15500001	15610000
AP008208_156	15600001	15710000
AP008208_157	15700001	15810000
AP008208_158	15800001	15910000
AP008208_159	15900001	16010000
AP008208_160	16000001	16110000
AP008208_161	16100001	16210000
AP008208_162	16200001	16310000
AP008208_163	16300001	16410000
AP008208_164	16400001	16510000
AP008208_165	16500001	16610000
AP008208_166	16600001	16710000
AP008208_167	16700001	16810000
AP008208_168	16800001	16910000
AP008208_169	16900001	17010000
AP008208_170	17000001	17110000
AP008208_171	17100001	17210000
AP008208_172	17200001	17310000

QY 333 -----AspLeuPheVala 337
 Db 80944 CATCTTTGAAAAATGTTAACTCCCTTTCTGTACACCTCCAGATGGAATTTCTTTTCAA 80885
 QY 337 laThrProSerAenIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuV 357
 Db 80884 AAACACGGAATGATGATCTCTCGAATGGTTGAAATTTTACAAATGTGACACTTGACATG 80825
 QY 357 alaAlaGlnAlaPheIleGlyLysLysProHISleuLeuGlnAAspLysGlnPheG 377
 Db 80824 TCCCTCAAGCAATTAATGATGAAAGCCAAATTCATCATGATGAGCAATCTTTTG 80765
 QY 377 lngInleuGlnGlnAlaLysValMetAlaMetGluIleProAlaMetLeuLysAspThrA 397
 Db 80764 AACCACTCCAGAAAGTTCAACATGAGTTCT-----CATTAATGCTTAATGATAT- 80718
 QY 397 rgaAenAntTpgIuIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyL 417
 Db 80717 -----GAGATGACCTTGCACTTGGAAGGGCAATTCGTCTATGCTAGTCGAG 80669
 QY 417 ysaValaArgGluCyAlaMetTTrpleuGlyLeuAspSerGluAspSerGlnTyrArgAsp 437
 Db 80668 ATGTTAGCAAGTCAGATGATGCTTGAAATGATATGATGATCTTCAACATACAGAGACC 80609
 QY 437 roAlaIleValaGluPheValleuGluAsnSerAen---ArgAspAspAsnAspAspLeuP 456
 Db 80608 CCAAAATTTCTAGAGTTATGTGACCAACTTTCAGCATCAGTGAAGAGAAATGATCTTTTC 80549
 QY 456 roGlyLeuCyAlaLeuGluGluThrTrpleuAlaGlyValaPheProArgPheArgA 476
 Db 80548 CAGGGCTGTCAAGCTTTTGAGACTTGCGCTTATCTTGAGTTTCTTCTAGAGAGCAGAG 80489
 QY 476 spThrLysAspLysLysPheLysLeuGlyAspTyrTyrAspAspPheValleuSert 496
 Db 80488 ATACTCGGGGCATGACGTTGAGACTTGGAATTAATTAATGATGATGATGATGATGATGAT 80429
 QY 496 ytrLeuGluArgValaGluValaGlnGlySerProLeuAlaAlaAlaAlaAlaAlaAla 516
 Db 80428 ACCTAGAAGAGATGAGAGGCTGCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 80369
 QY 516 rglIleGlyAlaGlu-----HisValLysAlaSerAlaMetGlnAlaL 530
 Db 80368 AACTGTGCTCAAGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 80309
 QY 530 eugLysValaPhePro---SerArgTyrThrAspArgAsnSerAlaGluProLysAspV 549
 Db 80308 TCAACAAAGTCTTTCATGATGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAG 80253
 QY 549 aIlgInGluThrValPheSerValaAspProValaGlyAAsnAAsnValaGlyArgAspGlyGluP 569
 Db 80252 -----AAAGATGGC---C 80243
 QY 569 roGlyValaPheIleAlaGluAlaValaArgProSerGluAsnPheGluThrAsnAspTyrA 589
 Db 80242 CTGGGGGATATCTT-----GAAAATTTTGAACAG----- 80214
 QY 589 laIleArgAlaGlyValaSerGluSerSerValaAspGluThrThrValaGluMetSerVala 609
 Db 80213 -----GAAAATGACACTG 80201
 QY 609 laAspMetLeuLysGluAlaSerValaLysIleLeuAlaAlaGlyValaAlaIleGlyLeu 629
 Db 80200 CTCAATGATTCGAAATATCGCTTGAAATATCTCTGCTGCGGCACTGCTTGTCACATGCT 80141
 QY 629 leSerLeuPheSerGlnLysTyrPheLeuLysSerSerSerSerPheGlnArgLysAspM 649
 Db 80140 TGGCAGATATTTGGGGCCAAATAT-----TTGCCTCTGTAAGAGGC 80102
 QY 649 etValSerSerMetGluSerAsp-----ValaIleThrIleGlySerValaArgAlaA 666
 Db 80101 CCTTCTCTGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80042

QY 666 spAspSerGlu----- 669
 Db 80041 CTGATGTGATGAAATATGCTTTTCATTGCTAAAGTAGTCAAAACAACATGGAATTTACTG 79982
 QY 669 ----- 669
 Db 79981 CTCACGTGTTTACTGGGATTCCTTGTGGCTGATTAACATATGTTCTTATATAGCT 79922
 QY 669 ----- 669
 Db 79921 TTGCATTATATGTTTAACAAAGCTATACATCAAAATATTTTGTCTCATATACACATG 79862
 QY 669 ----- 669
 Db 79861 CTACTATATCTGTAATGATATATCATAGAGAGTTAACTTTTGCATGATGAATGTTGAAG 79802
 QY 670 -----AlaLeuProArgMetAspAlaArgThrAlaGluA 681
 Db 79801 ATCTTGCACTAGATGAAGATCCAGTACATATTCCTGAATGATGAGGAGCTGGCAGAG 79742
 QY 681 snIleValSerLysTpgIuLysIleLysSerLeuAlaPheGlyProAspHISArgIleG 701
 Db 79741 ATATTTTGGCAAGTGGCAGAGATGATTAATCTTAAGGCTTTGGGACCAAGAACTTGGGTTG 79682
 QY 701 lueLeuPProGlu----- 705
 Db 79681 CATCATTTGCAAGAGGTGATGTTTAAAACTTTGACTTTTCATGCTTCTAGCTGGCCACTG 79622
 QY 705 ----- 705
 Db 79621 CTCTTACTATACCTAGCTTCTCTATAGAGATTAATTCAGATATACTCTCTATAGGA 79562
 QY 705 ----- 705
 Db 79561 TATATTTGTCAGCAATGTAACCTTTCAAGTTTGTGTAACATATCTTTCTTTGTTGACT 79502
 QY 705 ----- 705
 Db 79501 CACACATGCTTCAATTCGAATGCACTGAACAATTTTCAAGAACTATTGCACTGATGATA 79442
 QY 706 -----ValleuA 708
 Db 79441 TCCCGGTGATGAGGGGTCAATGCTGTGCTAGTATTTTCATGTTCTTATAGGCTCTTG 79382
 QY 708 spGlyArgMetLeuLysIleTpgThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuV 728
 Db 79381 ATGGCAACATGCTTAAGAGGTGATGATGACCGAGACCGAGATGAGCGTCAAGGGTGTG 79322
 QY 728 aLysAspTyrThrLeuLeuLysLeuSerValaAspSerValaThrValaSerAlaAspGlyT 748
 Db 79321 TCTGGAGATATACATATCCGATGACATGATGATGATGATGATGATGATGATGATGATGAT 79262
 QY 748 htrGAlaLeuValaGluAlaThrLeuGluGluSerLysLysSerLysSerLysSerLysSerLys 768
 Db 79261 GAGGAGCGACTGAGGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 79202
 QY 768 roGluAsnAAsnAlaThrAspValaArgThrTyrThrThrArgTyrGluValaPheTpgSerT 788
 Db 79201 CCAAGAAACATGATGATATATGACACAAATATCACTCCCGATGATGATGATGATGATGATGAT 79142
 QY 788 ys---SerGlyTpgLysIleThrGluGlySerValaLeuAlaSer 801
 Db 79141 AGCTAGAGAGGTGGAAGATACGGAAGAGCACTCTCAAGTGTG 79098
 RESULT 8
 AP004885 150462 bp DNA linear PLN 15-SEP-2004
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 DEFINITION PAC clone: P0575F10.
 ACCESSION AP004885
 VERSION AP004885.3 GI:41053009
 KEYWORDS Oryza sativa (japonica cultivar-group)

ORGANISM
Oryza sativa (japonica cultivar-group)

REFERENCE
AUTHORS
TITLE
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa niponbare (Ga3) genomic DNA, chromosome 2, PAC
clone: P0575F10
JOURNAL
TITLE
2
Published Only in Database (2002)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: sasaki@nias.affrc.go.jp URL: http://rpg.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Jan 21, 2004 this sequence version replaced gi:38142429.
Genes were predicted from the integrated results of the following:
GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), Genemark_hmm
(http://opal.biology.gatech.edu/Genemark/), Glimmer
(http://www.tigr.org/tdb/glimmer/glimmer_form.html), RiceHMM
(http://rpg.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), slim4
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologues of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from T7 to SP6 of the PAC clone.
This sequence of P0575F10 clone has an overlap with OJ1020.C02
(DBJ: AP004078) clone at 5' end and with P0482F12 (DBJ: AP005311)
clone at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rpg.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1. 150462
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Niponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0575F10"
/complement(4987..8112)
/gene="P0575F10.1"
/complement(join(4987..5412,5520..5679,5764..5914,
6096..6247,6738..6859,7864..8112))
/gene="P0575F10.1"
/note="supported by full-length cDNA(s): AK106553"
/complement(join(5003..5412,5520..5679,5764..5914,
6096..6247,6738..6859,7864..8084))
/gene="P0575F10.1"
/note="supported by full-length cDNA(s): AK073059"
/complement(join(5315..5412,5520..5679,5764..5914,
6096..6247,6738..6859,7864..7876))
/gene="P0575F10.1"

gene
/note="contains EST(s): AU070160(B61534)
contains full-length cDNA(s): AK106553,AK073059"
/codon_start=1
/product="putative eukaryotic translation initiation
factor 3, subunit 1 alpha, 35kDa"
/protein_id="BAD07941.1"
/db_xref="GI:41053010"
/translation="MEDWSDSEDFQAPVAPSAKAEPLKSMADDEVEDVKSWEDEE
EKRPPEVETAPKPSGKAQKQASSEVYQDEALDDPLETLROORLYEAD
FKSTTELFGKDGSEKSLDPIFKPSESDFPAVETLTKRTPREKSGWGLKNWR
LSMASLKGADAKDISSITAIINBKIAEKBAAGKKQAKKQILHTEKDDDFITG
KGNFDDDEYDFM"
complement(10186..13754)
/gene="P0575F10.2"
complement(join(10186..10479,10759..10892,11127..11991,
12089..12349,12515..12712,12843..13073,13455..13754))
/gene="P0575F10.2"
/note="start and end point are not identified"
complement(join(10186..10479,10759..10892,11127..11991,
12089..12349,12515..12712,12843..13073,13455..13754))
/gene="P0575F10.2"
/note="contains EST(s): AU058423(B51136), AU183658(B51136)"
/codon_start=1
/product="plastid division protein precursor"
/protein_id="BAD07942.1"
/db_xref="GI:41053011"
/translation="MGGFHLRLAPNSAPPAFSLPRPPRRPPPPSAACRAAR
MARLRPADPHLPTAASDPPSPAPAPAAAPSAFPPPLPPDAERSLPIQVDYKYL
GAPBFGDIDIRFEARLAKPOLYGTSTALGRPMVLIAHTLNNQNSRTYDRA
LSNREHFLTMDIAMDKEGALAVLTGQLLDRPKKFDVIVAMLAAYDLR
DMAASPPDVIGCEVLEEARLKLOEBGALAPDLISQIDETLEETFCVLELIL
PIIETHKKKQOEGLOGBRNLTWMSGRGITVGGGFRBAFNMBAFIRMISBOMDF
SDPISIPPEPPIYVAVLAHVAALGATISRRPPTPMADLPEOLKRNIGSHVAYNE
MDLALERAFCSLLVGVSKRMLGIDNESSPTDRPILFTYNSISSEBLLPGL
CKLETWLTFEVPERSDTRGMQFRGLDYDDDEVLSYLRMEGQASHLAAAAILK
LGAQATLALGTIVSNMIAQFNKFPLEIQDRAMENTKQPGVYLENFOENAPAD
SRNAALKTISGALFALAVIGAKYLPKRKPLAISSEHGVAVANSVDSTDPALD
DWHIIRPMADKLADIYRKQSIKRSALGSEHSAVLSQEVLDGNNMLKVMIDRAEIR
HPWFMYETLSDTYIDTISITSLDGRRAVETIDBAQLDTVTEPRNNDSIDTKTTRY
EMAFSKGKWKITGAVLKS"
complement(16401..17341)
/gene="P0575F10.3"
complement(join(16401..16739,16841..17341))
/gene="P0575F10.3"
/note="start and end point are not identified"
complement(join(16401..16739,16841..17341))
/gene="P0575F10.3"
/note="similar to Arabidopsis thaliana chromosome 3,
At3g61870"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07943.1"
/db_xref="GI:41053012"
/translation="NMAAPLAVAHITCSASNKNSPPSARQOQOTTTTATGSPAA
LPSILRTAAATAATTAALAPALAPALAAAGTLLVGLPVGAFATYLVIG
VIGMORRYVTTIDDEINELKKOLKPAALAAATPAAGVAGSSSSSPSPAPVPI
KIDELTBERRKLIKGSFRDRHFNAGSILILGCTBSYGALNTMFRGKLPFRPHLPA
GAATVLMMAAALVPMQKGNSTARSILTAALNVLFTWQIPGLSLVGVKVFET
TWP"
complement(18499..18942)
/gene="P0575F10.4"
complement(<18499..18942)
/gene="P0575F10.4"
/note="start and end point are not identified"
complement(18499..18942)
/gene="P0575F10.4"
/note="contains EST(s): AU184851(R10374), AU184850(R10374)"
/codon_start=1
/product="putative EF-hand Ca2+-binding protein CCD1"
/protein_id="BAD07944.1"
/db_xref="GI:41053013"
/translation="NMHHQOIRSTTAABQOQASAGGGGGEYEDLMPVNAIGAGSL
LSBLRAGFRLLADPARGATTAESTLRSSAASVILGIGGGGEMTVTEAAMAVREGDQGD

CDS

Gene: GAL5ABPCVIAWRLSPGIMGDAEGLMEALADELRLSLPPPPA"
21937..23552
/gene="P0575F10.5"
Join(<21937..22225,22326..22648,22736..23049,
23393..23552)
/gene="P0575F10.5"
/note="start and end point are not identified"
Join(21937..22225,22326..22648,22736..23049,23393..23552)
/gene="P0575F10.5"
/note="similar to Arabidopsis thaliana chromosome 5,
At5g24330"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07945.1"
/db_xref="GI:41053014"
/translation="MGPRATPLRRRTARRAPATRAEGSGDGDVVRCAGCGGSA
AELLDCDCGRGHIIFCLRPILPRVPAGDFPCSPSHSKSMAAKKPKPVLVQ
TKIVDFKIQGRPAALAAAESEEGKRRKVGIRLVSKRRKLPERSPDPAK
LRMASLATALTATGAVFSNELTYVPGMARANRALESQGQVLEKEVDLNLCK
RMAAGMPEPLVVDVEGTVADRFIDLTITVGVVDVYLFRHDDSDMWT
LSAATPSRLVIPCGRSNLAFINGINHTPDGRKKQMLKCRVPDVGECVLLVA
NRDISGERLYIYNGSEHEYPTHHPY"
CDS
Complement (Join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/gene="P0575F10.6-1"
/note="supported by full-length cDNA(s): AK121372"
Complement (Join(24491..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28147))
/gene="P0575F10.6-1"
/note="supported by full-length cDNA(s): AK062255"
Complement (24563..28147)
/gene="P0575F10.6-2"
Complement (24563..28147)
/gene="P0575F10.6-2"
/note="contains full-length cDNA(s): AK065148
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
Complement (Join(24787..24817,24913..25266,25398..25633,
25720..25832,27054..27185))
/gene="P0575F10.6-1"
/note="contains EST(s): AU031017(E60571), AU093466(E60571)
contains full-length cDNA(s): AK121372, AK062255"
/codon_start=1
/product="putative arginine/serine-rich splicing factor"
Alignment Scores:
Pred. No.: 1,79e-86 Length: 150462
Score: 154050 Matches: 398
Percent Similarity: 41.7% Conservative: 118
Best Local Similarity: 32.2% Mismatches: 215
Query Match: 37.9% Indels: 505
DB: 15 Gaps: 20
US-10-600-070b-2 (1-801) x AP004885 (1-150462)
QY 12 SerProhGlnLeuCyAArgLeuProAlaThrThryLeuAlaGArgSerHisAAsn 31
Db 13715 GCGCAATTCCTCTCTCCCTCCCTCGCGCGCGCGCGCGCGCGCGCGCT 13656
QY 32 ThSerThrThrIleCyS---SerAlaSerIleThrAlaAspArgLeuLeuSerAspHe 50
Db 13655 CACCCCTCCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCGCACTTC 13596
QY 51 AsnPhe-----ThSerAspSerSerSerSerSerPheAlaThAlaThr 65
Db 13595 CACTCTCTCCCAACCGCGCGCGCTCCGACCGCGCGCTCCCGCGCGCGCGCGCC 13536
QY 66 ThrThAlaThrLeuValSerLeuProProSerIleAspArgProGluArgHisValPro 85
Db 13535 GCGCGCTCCGCGCTCCCTTCCTCCGCTTCCCGACGCGCGCGCGCGCGCTCCCTCCG 13476

QY 86 IleProIleAspPheTy----- 91
Db 13475 CTCGAAGTGATTTCTACAAAGGAGGAGCCCGCGAATCTGTCGCGTTGTTCATCATTT 13416
QY 91 ----- 91
Db 13415 GGAAGCTCGAATATCATCCCGTAGATMAAAGGCGATGATCTTGCCTTCTGTGTAG 13356
QY 91 ----- 91
Db 13355 GTGATGATACATTTGATCAGTCTTAATCTGTTGAAAGCTTATTCACCTGCAAGCTTTTGC 13296
QY 91 ----- 91
Db 13295 TAGGTTACTGATCCCTCAGTTCATGAAATCGAGTTATTTGCTATTTCCGTAGTCACT 13236
QY 91 ----- 91
Db 13235 TCAGTAGTTATTAAGTCACTAATCGAATCGAATCATCGCATTTGTATCCACTGATAC 13176
QY 91 ----- 91
Db 13175 TATGATACGAATTATGCTTATCATGTGTGTGTGGGCTAGAGATCATGCTTTT 13116
QY 92 ----- GlnValLeuGlyValGlnThr 98
Db 13115 TGACATATTCGAATGTTTGTGTGAATCTGCTTGCCTTGCAGGTTCTAGGGGAGAGACCA 13056
QY 99 HisPheLeuThrAspGlyIleArgAlaAlaPheGluAlaArgValSerIleProProGln 118
Db 13055 CATTTCCTTGGCCATGATGATCAGAGAGGGGCTTCGAGCAGCGATCCAAAGCCGCGCAG 12996
QY 119 PheGlyPheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCysGlu 138
Db 12995 TATGGCTACGACGACGAGATGCTTGTGTGCTGTCGACAAATGCTGAGATGCCCATGAC 12936
QY 139 ThrLeuSerAsnProArgSerArgArgArgGluTyArgGlnGlyLeuLeuAspAspGlnGlu 158
Db 12935 ACTCTGATGACGACGAACTCCGCACTGATGATGATGATGATGATGATGATGATGATGAT 12876
QY 159 AlaThrValIleThrAspValProTrpAspIle----- 169
Db 12875 GAAGCTCTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12816
QY 169 ----- 169
Db 12815 GTGTTGGATCTTGTCAATAGACATAGTGTATATCTCTCTGCTTATATATATG 12756
QY 170 ----- ValProGlyAlaLeuCyValLeuGlnGlyGlyGluThrG1 184
Db 12755 GTGCTTTGTCAGACAGGTTCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 12696
QY 184 uIleValLeuArgValGlyGlyAlaLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 204
Db 12695 TGCTGTGCTTGTATCTGAGAACAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 12636
QY 204 naSPValIleValIleValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLe 224
Db 12635 GGAAGT 12576
QY 224 uAspProProAspPheIleThrGlyTyArgIlePheValGlyGlyAlaLeuLeuLeuGly 244
Db 12575 AAGCCCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12516
QY 244 n----- 244
Db 12515 GTGAACTGTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 12456
QY 244 ----- 244
Db 12455 AGATTAAATACCTGTGAAGATGTTGTGGACACATGGGCTATATTTTAAATAGCA 12396

QY 245 -----GlutInGlyAlaSe 249
 DB 12395 ATGCGTATTCCTTGAAGCATATTAAATAGCACTTCACTGATGAGAAAGAGCAAG 12336
 QY 249 rSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGlnGluIleThrProAr 269
 DB 12335 CAATCTGCAAGCTGATCTGCTTTCACAGATGATGAACCTCTCGAGAGATTAACCTCG 12276
 QY 269 gTyValLeuGlnLeuLeuGlyLeuProLeuGlyAspAspTyValAlaAlaGluLeuAs 289
 DB 12275 CTGTGATTTGAGCTTCTCTCCCTTCTATTTGACACAGCATATAGAAGCCGACAGA 12216
 QY 289 ngIlyLeuSerGlyValArgAsnIleLeuThrPseValGlyGlyGlyAlaSerAlaLe 309
 DB 12215 AGGCTTCAAGGTCGCGGAAACATTTTGTGAGCGCTTGGCAGAGAGATATGCTTACCT 12156
 QY 309 uValGlyGlyLeuThrArgGluTyPheMetAsnGluAlaPheLeuArgMetThrAlaAl 329
 DB 12155 TGGAGGAGATTTTCTGTCGTCGATGAGCTTCATGAACGAGCTTTTGTGAGATGACATCAAT 12096
 QY 329 agIuGlnVal----- 332
 DB 12095 TGAACAGTTTCATCATGTTGTTTATGAGACAGCTGCTTATTTATTCAGACAGACA 12036
 QY 333 -----AspLeuPheValA 337
 DB 12035 CATCTTGAACAACTGTTAACTCCCTTCTGTTACCTCCAGATGAGATTTCTTTTCAA 11976
 QY 337 laThrProSerAsnIleProAlaGluSerPheGluValTyArgIuValAlaLeuAlaLeu 357
 DB 11975 AAACACCGAATAGCATTCCTCCGTAATGTTGAAATTTTCAATGATGACCTTGACATG 11916
 QY 357 alAlaGlnAlaPheIleGlyValTyLeuProHISLeuLeuGlnAspAlaAspTyLeuPheG 377
 DB 11915 TCCTCTAAGCAATATATAGTAAAGGCCACATTCATCATGATGCGGATGATCTTTTGG 11856
 QY 377 lngInLeuGlnGlnAlaIleValMetAlaMetGluIleProAlaMetLeuTyAspThrA 397
 DB 11855 AACAACTCCGAAAGTTCAACATAGCTTCT-----CATTATGCTTATATATAT- 11809
 QY 397 rghAsnMetTPGluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyL 417
 DB 11808 -----GAGATGACCTTGCATTTGAAAGGCGATTCCTGCTCATCTTGACTCGAG 11760
 QY 417 ySValAspGluCyArgMetCTPleuGlyLeuAspSerGluAspSerGlnTyArgAsp 437
 DB 11759 ATGTAGCAAGTGCAGATGTGCTTGAATGTATATGATGATTCACCAATACAGAACCC 11700
 QY 437 roAlaIleValAlaGluPheValLeuGluAsnSerAsn-----ArgAspAspAspAspAspLeu 456
 DB 11699 CCAAAATTTCTAGAGTTTATTGTGACCAACTTACATGATGATGAGAAAGATGATCTTCTC 11640
 QY 456 roGlyLeuCyAlaLeuGlnGluIleThrTPleuAlaGlyValAlaPheProArgPheArgA 476
 DB 11639 CAGGGCTGTCCAGACTTTTGAAGCTTGGCTTATCTTGAAGCTTTTCTTGAAGAGCAGAG 11580
 QY 476 spThrIleAspTyLeuPheLeuGlnGlyAspTyTyArgAspAspPheMetValLeuSer 496
 DB 11579 ATACTCGGGCAGATGCTTGAAGCTTGAAGTACTGATGATGATGATGATGATGATGAT 11520
 QY 496 yTrLeuGluArgValGluValAlaGlnGlySerProLeuAlaAlaAlaAlaAlaAlaAla 516
 DB 11519 ACCCTGAAGAGATGAGAGGCTGCTGCTTCTCATTTGCTGCTGCTGCTGCTGCTGCTGCT 11460
 QY 516 rgIleGlyAlaGlu-----HisValIleValAlaSerAlaMetGlnAlaL 530
 DB 11459 AACTGTGCTCAACAGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11400
 QY 530 euGlnIleValPhePro-----SerArgTyThrAspAspAspAspAspAlaGluProTyAsp 549
 DB 11399 TCAACAGAGTTTTCATTGATAGAACATTTAGACAGCTGCGCATGAGAAATATCT- 11344
 QY 549 alGlnGluThrValPheSerValAspProValGlyAsnAsnValGlyArgAspGlyGluP 569

DB 11343 ----- 11334
 QY 569 roGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAspTyArgA 589
 DB 11333 CTGGGGATATCTT-----GAAAATTTTGACAG- 11305
 QY 589 laIleArgAlaGlyValSerGluSerSerValAspGluThrThrValGluMetSerValA 609
 DB 11304 -----GAAAATGCACCTG 11292
 QY 609 laAspMetLeuTyGluAlaSerValIleValAlaAlaGlyValAlaAlaIleGlyLeuI 629
 DB 11291 CTGATGATTTGAGAAATGCGCGCTTGAAGTTATCTGCTGCGGCACTGTTGCACTGT 11232
 QY 629 leSerLeuPheSerGlnTyTyPheLeuTySerSerSerSerPheGlnArgTyAspM 649
 DB 11231 TGGCAGTAAATTTGGCGCCAAATAT-----TTGCTCTGTAAGAGGC 11193
 QY 649 etValSerSerMetGluSerAsp-----ValAlaThrIleGlySerValArgAlaA 666
 DB 11192 CCTTCTGCTATTTAGAGTGAAGCATGATCTGTGCGATTTGCTTAATAGTGCACCTTA 11133
 QY 666 spAspSerGlu----- 669
 DB 11132 CTGATGTGAGTAAATGCTTTTCAATTGCTAAAGTAGTCAACACACATGGAATTTACTG 11073
 QY 669 ----- 669
 DB 11072 CTCACGTGTTACTGGGTATCCTTGTGCTGATTAACATGATGCTTGTGATCTTATAGCT 11013
 QY 669 ----- 669
 DB 11012 TTGCATTAATGTTTAAACAAAGCTAATCAATCAAAATTTTGTCTCATACATACAAATG 10953
 QY 669 ----- 669
 DB 10952 CTACTAATATTCGTATATGATATCATAGAGTTAACTTTGCATGATGATGATGATGAG 10893
 QY 670 -----AlaLeuProArgMetAspAlaArgThrAlaGluA 681
 DB 10892 ATCTGCACTAGATGAAGATCCAGTACATATTTCTTGAATGAGTGCACACCTGCGAGAG 10833
 QY 681 snIleValSerIleTPGlnIleIleSerLeuAlaPheGlyProAspHisArgIleG 701
 DB 10832 ATATGTTGCGAAGTGCAGATGATCAATCTAAGCTTGGAGCAGAACTTGGGTG 10773
 QY 701 lMetLeuProGlu----- 705
 DB 10772 CATCATTTGCAAGAGGTGATTTAAACCTTTGACTTTTCAATGCTTCTTACCTGCGCAC 10713
 QY 705 ----- 705
 DB 10712 CTCTTACTTATCTTACCTGCTTCTTATAGAGATTAATTCAGATTAATCTCTTATAGGA 10653
 QY 705 ----- 705
 DB 10652 TATATTTGTCAGCAATGTACTTTTCAAGTTTGTTCACATATCTTTCTTGTGACT 10593
 QY 705 ----- 705
 DB 10592 CACACATGCTTCAATGCACTGAACAAATTTTCAAGAACTTATGCCACTGCAATGATA 10533
 QY 706 -----ValIleuA 708
 DB 10532 TCCCGTATGATGCGGCTCAAAATGCTGTAATTTTTCATGTTCTTATAGATTTCTTG 10473
 QY 708 spGlyArgMetLeuTyIleTyThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeu 728
 DB 10472 ATGCAACATGCTTAAAGTGTGACTGACCGAGCGGAGATTTGAGCGCATGAGGTGCT 10413
 QY 728 alTyAspTyThrLeuLeuTyLeuSerValAspSerValThrValSerAlaAspGlyTy 748

Db 10412 TCTGGAGTATACATCCGATGTACGATTGATGACATCACTATCTCCCTAGATGCTC 10353
Qy 748 hraralaleuValaIGluAlaThrleugluGluSerAlaCysLeuSerAspLeuValHis 768
Db 10352 GACGAGGAGCTGTGAGGCTTACATGATGAGGAGCCAACTTACGATGTTACTAGC 10293
Qy 768 rogluenaAenAlaThrAspValArgThrTyrThrThrArgTyrGluValPheTrpSerL 768
Db 10292 CCAGAAACATGATTCATGATGACACAAATACTACCTCCGTTAGATGAGGCTTCCTCA 10233
Qy 788 yS---SerGlyTyrPlyValIleThrGluGlySerValLeuAlaSer 801
Db 10232 AGCTAGAGGGGTGAGATTAACGAGAGGACGCTCTCAAGTCTG 10189
RESULT 9
AC158210
LOCUS Medicago truncatula clone mch2-155019, WORKING DRAFT SEQUENCE, 4
DEFINITION Medicago truncatula clone mch2-155019, WORKING DRAFT SEQUENCE, 4
AC158210 133779 bp DNA linear HTG 22-JUN-2005
AC158210
AC158210.11 GI:71061528
HTG: HTG PHASE2; HTGS DRAFT
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
REFERENCE 1 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Medicago truncatula BAC Clone mch2-155019
JOURNAL Unpublished
2 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 133779)
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Jul 22, 2005 this sequence version replaced gi:68342147.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKMOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 43719: contig of 43719 bp in length
* 43720 43819: gap of unknown length
* 43820 83929: contig of 40110 bp in length
* 83930 84029: gap of unknown length
* 84030 106272: contig of 22243 bp in length
* 106273 106372: gap of unknown length
* 106373 133779: contig of 27407 bp in length.
* Location/Qualifiers
1..133779
/organism="Medicago truncatula"
/mol_type="genomic DNA"

/db_xref="taxon:3880"
/clone="mch2-155019"
/clone_lib="Medicago truncatula BAC library H2"
43720..43819
/estimated_length=unknown
83930..84029
/estimated_length=unknown
106273..106372
/estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.: 1,82e-67 Length: 133779
Score: 1241.00 Matches: 319
Percent Similarity: 33.9% Conservative: 63
Best Local Similarity: 28.3% Mismatches: 87
Query Match: 30.5% Indels: 659
DB: 14 Gaps: 9
US-10-600-070B-2 (1-801) x AC158210 (1-133779)
Qy 7 ValGlyIleGlyLeuSerProPheGluLeuCysArgLeuProAlaThrThrLeu 26
Db 130423 ATGGGCTCAGCTCTTCCACTTCTCCG-----CGTTCACAGCTTCTCCAAAACCT 130476
Qy 27 ArgArgSerHisAsnThrSerThrThrIleCysSerAlaSerIleTyrPalaAspArgLeu 46
Db 130477 AACCGTCCAT-----TCCTCGCGGCTCCGCCACACAGTAATAGGGGAGCACTC 130530
Qy 47 LeuSerAspPheAsnPheThrSerAspSerSerSerSerPheAlaThrAlaThrThr 66
Db 130531 ATTTCCGATTTCCAAATTCCTCGGCGACACCTCTCTTCTCTCCACACACACCTCCGCC 130590
Qy 67 ThrAlaThrLeuVal---SerLeuProProSerIleAspArgProGluArgHisValPro 85
Db 130591 ACAGTCACTCTCACTCTCTTCACTCTCTCCGATA-----GAACCCACGCTGCA 130641
Qy 86 IleProIleAspPheTyrGluValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIle 105
Db 130642 CTCCTCTCGACCTGTACAAATATCTCGGGCGGAAACGATTTCTCGGTGATGATAT 130701
Qy 106 ArgArgAlaPheGluAlaArgValSerIleProProGluPheGlyPheSerAspAspAla 125
Db 130702 CGAGAGCTTATGAACGAATCTCGAAGCCCTCCAGATCTCTTCAGTATGAAGCT 130761
Qy 126 LeuIleSerArgArgGlnIleLeuGlnAlaCysGluThrLeuSerAsnProArgSer 145
Db 130762 TTGATTAGTCGTCTGATGATTTCAAGCTGTGTAACCTTACGTATCTGCTTCT 130821
Qy 146 ArgArgGlyTyrAsnGluGlyLeuLeuAsp-----AspGluGluAlaThrVal 161
Db 130822 AGAAGAGATTAATATCAAGCCCTCGACAGATGAAGAAAGCAAGATGAGGAATCTTCAT 130861
Qy 162 IleThrAspValProTrpAsp-Lys----- 169
Db 130862 CTCACGAAATCCCTTTCGACAAAGTAAGCTTCAATTCAATTCAATTCAATTCAATT 130941
Qy 169 ----- 169
Db 130942 GAGTTTGTCTTCATTAATAAATGATTAAATGTTGTAATGCAATGATATAGATTCA 131001
Qy 169 ----- 169
Db 131002 TTTACTAATAACACTTTTATTAGTGAAGTACTGATTAGTAAGATTAAGATTAATAAAT 131061
Qy 169 ----- 169
Db 131062 AGTTTAAGACATGTTTTCGAAAAACAGTGAATAGTTGATTAACAGACTTATAGTGATGA 131121
Qy 169 ----- 169
Db 131122 GCTTATGAAGATTAAGCTAGATGATTAATTTGTAGTGTGTTGTAACCTTAGCGGTTAAAC 131181

QY 169 ----- 169
 Db 131182 TAGCTTATATATATGAATGATGAAAAATATATATTTTAATTAATATATATTTT 131241
 QY 169 ----- 169
 Db 131242 TCAATTTAGATTATAGGGTAAATGAGAGAAAAATGTAACTTAACTACTTGAG 131301
 QY 169 ----- 169
 Db 131302 TTAGCTTATAAAGAACTTTAAGCTAAGTAACTAAGCTTATAGCTGCGCAAAATGAC 131361
 QY 169 ----- 169
 Db 131362 AATTGTCTTTTTCATGACATAGCATTAAGCTAAGAAATGTGTCTGGCAACAATAC 131421
 QY 169 ----- 169
 Db 131422 CTTGAGTGTCACTTTGAGCTCGGCTTCAAAAAGAAATGATTTATGTGTGTGATTT 131481
 QY 169 ----- 169
 Db 131482 CAATGATAAATAGATTTCATCTAAACAATCCTTATTAATGTGAGATCTAGAGA 131541
 QY 169 ----- 169
 Db 131542 AATTAAGTACAAATTAAATAGTTAAATTTAGTTGGGAAAAAATAGTTAATGTTTCATTAG 131601
 QY 169 ----- 169
 Db 131602 TGTTCTAAAGTCACTATTATTTTGGATAGAAAAATTTGCCAATGAGACAATTATACG 131661
 QY 169 ----- 169
 Db 131662 AGACAGCGCATATTTGTTAATGTTAATGCAATGATCTTTTATTTTGTGTTTTGTA 131721
 QY 170 -ValProGlyAlaLeuCySerValLeuGlnGlyGlyGluThrGluLeuLeuArgVa 189
 Db 131722 GGTTCCTGAGCTCTGTGCTGTGTTGCAAGAGCTGAGAGCGAGGTGGTCTTGCAAT 131781
 QY 189 IGIYGLUALALEULYSGLUARGLEUPROLYSSEPHLYSGLNASPYALILEUVA 209
 Db 131782 TGGAGGGGGTTTACTGAGAGAGAGGTTACCGAAGATGTTTAAGCAAGATGTGTGGTGC 131841
 QY 209 IMEALALEUALAPHEULASPYALSERARGASPYALMECALALEUSAPROPHAPPH 229
 Db 131842 TATGGCGCTTGCAATATGTTGACGTTTCTAGGAGTGTATGCTTGTCCCGCCAGATTT 131901
 QY 229 eIlethrgLyTyrgLupheValIGluGluAlaLeuLysLeuGln----- 244
 Db 131902 CATTTGTTGCTTGTGAGATGCTGGAAAGGCAITGAACTTTTGACAGGTGAGAGTAAATAC 131961
 QY 244 ----- 244
 Db 131962 CTAGTTTTCAGGCTATAAATATGTATTTACTTACAGTATGCTGTCTGGAATTTAGCT 132021
 QY 244 ----- 244
 Db 132022 TCTAGATTAAGAAATACATTAATATGTTTATAGATAATTTGATTGATTAACAAGTAA 132081
 QY 244 ----- 244
 Db 132082 AATGAGAGATTTTGGAGAAATTTGAAATGTATAGGCTAATTTCAACTGTGTAGAGAAC 132141
 QY 244 ----- 244
 Db 132142 AATTGTTTGAATAATGATGTCTATTAACAACCTTAATATTTTAATTTTGATAAGTAT 132201
 QY 244 ----- 244
 Db 132202 TGCAGACAGGTATATCATCTTGACTCTAGAGTACTAGAAATTAATACCCCGTAGACATA 132261
 QY 244 ----- 244

Db 132262 GAACATAGCCAGTGGCCCTTTTTCGCAATTGAAGTATTGAGAAATGCACTGATTAGACA 132321
 QY 244 ----- 244
 Db 132322 AACATATAGCTTTTTTTTATGTGAGCCATCATATGATTAAGATTTGTTACCGAAT 132381
 QY 244 ----- 244
 Db 132382 GAGTACAAAATAAAGAAAAAGTATTAATCTATGATGATGTTCAATGTAAAAAGTAAATA 132441
 QY 244 ----- 244
 Db 132442 TCTACCAAGTATTCAACCAATGTAAATTAACCATGCACTTTTACAAAATCATTTCTT 132501
 QY 244 ----- 244
 Db 132502 TAACTTGAATGATTAAGCCATATATGATGTATATTTTATTAACAATTTAT 132561
 QY 244 ----- 244
 Db 132562 AATGTTAATGAATGTCCGTGATTACTGTGTAGCAGAGCTCGTTACAAATTTGTGTAA 132621
 QY 244 ----- 244
 Db 132622 TCTTTCATGTCATTATTTGTAAGTCTTCTCTAATATGATGTAATCTTACTATCA 132681
 QY 244 ----- 244
 Db 132682 CTATTCGCACAAAATGATATGTGTGTCTGTTGTATATGATGTGGAATTAATAGTA 132741
 QY 244 ----- 244
 Db 132742 CTTGGGTAAATTTTGTATATTTGTCAAGTATAGTATCTTTTAACTTACTACTGT 132801
 QY 244 ----- 244
 Db 132802 GCATTTCCAAAATCTACACAAATATATTTTGTGATGATATTAATGTGATTAACAAATG 132861
 QY 244 ----- 244
 Db 132862 TGTATATGCTGAAGACTCAATTTGCAATGCTGTGCAATATGCTTCTGATGACTCACT 132921
 QY 245 -----GluGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspG 261
 Db 132922 GATCATGTAGGAAGAAAGGGCAAGCAGCTTAGACCGGATTTACAAACACAAATGATAGA 132981
 QY 261 uThrLeuGluGluIleThrProArgTyValLeuGluLeuGlyLeuProLeuGlyAs 281
 Db 132982 GACACTTGAAGAGATTAACCCCACTGTGTGTATTAAGAACTTTAGCCTTGCTTGATGA 133041
 QY 281 PAEPYRALAALALYARGLYUANGLYLEUSERGLYVALARGASNILLEUTRPSERVA 301
 Db 133042 TGAACATCGAGCGCGAGAGAAAGACAGGTCTCAAGGTGTCCGCAACATCTGTGGGCAGT 133101
 QY 301 IGIYGLYGLYGLYAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheLeuGln 321
 Db 133102 TGGAGGTGAGGAGCAGACCAATTTCTGTGAGATTTTCAACGTCGAGGATTTCAATGAATGA 133161
 QY 321 uAlaPheLeuArgMetThrAlaIleGluGlnValAsp----- 333
 Db 133162 GGCATTCCTGCAATATGAAGCTCCGAAACAGTGT- CATGAATATATTTGATGATGTT 133220
 QY 333 ----- 333
 Db 133221 TTATTTTAATTTCTTCGCCCCCTTAGTTTGTGTTATGAAGGTTTACATTTATCT 133280
 QY 334 -----LeuPheValAlaThr 338
 Db 133281 CTTTCATTAATCTTTTATCTCCAAATCTTTTATTAAGTCAACTTTTGTAGCAAC 133340
 QY 338 rProSerAsnIleProIleGluSerPheGluValTyrgLuvAlaIleuAlaLeuValAla 358
 Db 338 rProSerAsnIleProIleGluSerPheGluValTyrgLuvAlaIleuAlaLeuValAla 358

DB	133341	ACCTGACCAATATTCACGCTGAAAGTTTGAGGCGCTATGGGGTTGACCTTGACATGGTTGC	133400
QY	358	AGlnAaPheIleGlyLyLeValProHIsleuEngInaAPaAaPlyGlnPheGIneI	378
Db	133401	ACAAAGCCTTGTAGGTAAAGACCATCTTATCCAAAGATGTGATTAATTTATTCATCA	133460
QY	378	nLeuGInGInaIaLyValMetAlaMetGulIeProAlaMetLeuTyraSP-----Th	396
Db	133461	ACTTCACAACTTAAGTAAAGTAACAAATATGAGAAATGCTCCCTCTGTTTATCTCCATGGA	133520
QY	396	rArgAaenAentTpsGulIeAaSPheGlyLeuGInaTArgGlyLeuCysAlaLeuLeuIleG	416
Db	133521	CATGGAAGAAGAGAAAGTGCATTTTTCATTTAGAAAGGGCTGTGTGCACTGCTTGTGG	133580
QY	416	LyValaAaPgluCyAaRmeTTrPleuGlyLeuAaSPSerGluAaSPSerGInTyraRga	436
Db	133581	GGAGCTGTATCATATGTCATGTCATGTTGGGTCCTCGATAGAACAGCTCACCTTATCGAAA	133640
QY	436	nProAlaIleValGluPheValLeuGluAaenSeraAaRgaAaSPaAaSPaAaSPleuPr	456
Db	133641	CCCACTATTAATAGACTTATTAATGAGAAAACGCAAAAGGTGAAGAAAGACGTATCTTCC	133700
QY	456	oGlyLeuCyLySLeuLeuGluThTrPleuAlaGlyValaIValPheProAaRgPheAaRga	476
Db	133701	TGGACTCTTAATTTGTTGGAGACATGCTTGATGAGAGCTGCTTTCCCTTAGGTTTAGAGA	133760
QY	476	PThrLyAaPlyLySPha 482	
Db	133761	TACTAAAGACAACTTT 133779	
RESULT 10			
LOCUS	AC155597	184752 bp	DNA linear HTG 25-JAN-2005
DEFINITION	Zea mays strain B73 clone ZMMBc0216X08, ** SEQUENCING IN PROGRESS		
ACCESSION	AC155597.2	GI:58082456	
VERSION	AC155597.2	GI:58082456	
KEYWORDS	HTG; HTGS - PHASE1.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE			
AUTHORS	1 (bases 1 to 184752)		
	Chan, A. P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,		
	Uterback, T. R., Feldblum, T. V., Rabinowicz, P., Frazer, C. M.,		
	Schubert, K., Sam Miguel, P., Ma, J., Pontaroli, A. C., Rohlfing, T.,		
	Budiman, A., Bedell, J. A., Lakey, N., Barbazuk, B., Bennetzen, J., and		
	Quackenbush, J.		
TITLE	Consortium for Maize Genomics - BAC skim sequencing and assembly		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 184752)		
AUTHORS	Chan, A. P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,		
	Uterback, T. R., Feldblum, T. V., Rabinowicz, P., Frazer, C. M.,		
	Schubert, K., Sam Miguel, P., Ma, J., Pontaroli, A. C., Rohlfing, T.,		
	Budiman, A., Bedell, J. A., Lakey, N., Barbazuk, B., Bennetzen, J., and		
	Quackenbush, J.		
TITLE	Direct Submision		
JOURNAL	Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),		
REFERENCE	9712 Medical Center Dr, Rockville, MD 20850		
AUTHORS	3 (bases 1 to 184752)		
	Chan, A. P., Petrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,		
	Uterback, T. R., Feldblum, T. V., Rabinowicz, P., Frazer, C. M.,		
	Schubert, K., Sam Miguel, P., Ma, J., Pontaroli, A. C., Rohlfing, T.,		
	Budiman, A., Bedell, J. A., Lakey, N., Barbazuk, B., Bennetzen, J., and		
	Quackenbush, J.		
TITLE	Direct Submision		
JOURNAL	Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),		
REFERENCE	9712 Medical Center Dr, Rockville, MD 20850		
AUTHORS	On Jan 25, 2005 this sequence version replaced gi:57863118.		
COMMENT	----- Trace submision		
	Seq. name: TIGR		
	Seq. lib id: ZG30		

```

----- Project information
Web site: http://www.tigr.org/tgi/maize/
Contact: maize@tigr.org
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown..
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2210: contig of 2210 bp in length
* 2211 9941: contig of 7631 bp in length
* 9942 10041: gap of unknown length
* 10042 16745: contig of 6704 bp in length
* 16746 16845: gap of unknown length
* 16846 32394: contig of 15549 bp in length
* 32395 32484: gap of unknown length
* 32484 35494: contig of 3000 bp in length
* 35495 35594: gap of unknown length
* 35595 40051: contig of 4457 bp in length
* 40052 40151: gap of unknown length
* 40152 41379: contig of 1228 bp in length
* 41380 41479: gap of unknown length
* 41480 44517: contig of 3038 bp in length
* 44518 44617: gap of unknown length
* 44618 45564: contig of 947 bp in length
* 45565 45664: gap of unknown length
* 45665 49440: contig of 3776 bp in length
* 49441 49540: gap of unknown length
* 49541 52698: contig of 3158 bp in length
* 52699 52798: gap of unknown length
* 52799 55964: contig of 3166 bp in length
* 55965 56064: gap of unknown length
* 56065 59236: contig of 3172 bp in length
* 59237 59336: gap of unknown length
* 59337 70052: contig of 10716 bp in length
* 70053 70152: gap of unknown length
* 70153 84603: contig of 14451 bp in length
* 84604 84703: gap of unknown length
* 84704 86418: contig of 1715 bp in length
* 86419 86518: gap of unknown length
* 86519 88069: contig of 1551 bp in length
* 88070 88169: gap of unknown length
* 88170 89661: contig of 1492 bp in length
* 89662 89761: gap of unknown length
* 89762 91357: contig of 1596 bp in length
* 91358 91457: gap of unknown length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 169958: gap of unknown length
* 169959 176100: contig of 6142 bp in length
* 176101 176200: gap of unknown length
* 176201 183079: contig of 6879 bp in length
* 183080 183179: gap of unknown length
* 183180 184752: contig of 1573 bp in length.
*
* Location/Qualifiers
*
* 1..184752
*   /organism="Zea mays"
*   /mol_type="genomic DNA"
*   /strain="B73"
*   /db_xref="taxon:4577"
*   /clone="ZM85BC0216K08"
*   2211..2310
*   /estimated_length=unknown
*   9942..10041

```



```

gap      /estimated_length=unknown
16746..16845
gap      /estimated_length=unknown
32395..32494
gap      /estimated_length=unknown
35495..35594
gap      /estimated_length=unknown
40052..40151
gap      /estimated_length=unknown
41380..41479
gap      /estimated_length=unknown
44518..44617
gap      /estimated_length=unknown
45565..45664
gap      /estimated_length=unknown
49441..49540
gap      /estimated_length=unknown
52699..52798
gap      /estimated_length=unknown
55965..56064
gap      /estimated_length=unknown
59237..59336
gap      /estimated_length=unknown
70053..70152
gap      /estimated_length=unknown
84604..84703
gap      /estimated_length=unknown
86419..86518
gap      /estimated_length=unknown
88070..88169
gap      /estimated_length=unknown
89662..89761
gap      /estimated_length=unknown
91358..91457
gap      /estimated_length=unknown
132728..132827
gap      /estimated_length=unknown
137469..137568
gap      /estimated_length=unknown
163844..163943
gap      /estimated_length=unknown
169859..169958
gap      /estimated_length=unknown
176101..176200
gap      /estimated_length=unknown
183080..183179
gap      /estimated_length=unknown

```

ORIGIN

Alignment Scores:

```

Pred. No.:      4.41e-66      Length:      184752
Score:          1222.00      Matches:      361
Percent Similarity: 32.0%      Conservative: 116
Best Local Similarity: 24.2%      Mismatches: 183
Query Match:      30.1%      Indels:      834
DB:              14          Gaps:      17

```

US-10-600-070b-2 (1-801) x AC155597 (1-184752)

```

QY      90  PheTyrGlnValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIleArgAlaPhe 109
      107759 TATGCTCAAGATTCTGTGGCGAGCCACATTCTTACGCGATGCGAGGCGCTTC 107818
      110 GluAlaArgValSerLysProProGlnPheGlyPheSerAspAlaLeuIleSerArg 129
      107819 GAGTCGCGAGTACGTAAGCCACCTCAGTATGGGTACAGCACAGAAAGCTTGTGGCGCA 107878
      130 ArgGlnIleLeuGlnAlaGlySerGlnThrLeuSerAsnProArgSerArgArgGlnTyr 149
      107879 CGGCAAAATGCTGCAGATTGCCAGATGATCTCTCAAAACGAGAGCTGCGCAGCAATAC 107938
      150 AsnGlnIleLeuLeuAspAspGlnGluAlaThrValIleThrAspValProTyrAspLys 169
      169

```

```

Db      107939 GACCGTCGCTTCCGAGACCGTGATGCCGCACTACCATGATGTTCGCGATTAAG 107998
QY      170 Val----- 170
Db      107999 GT-GAGGTATGGACACAAACGTGCCAAATGCTATCTGTTTCTTGTAATCTGTA 108057
QY      171 -----Pro 171
Db      108058 AATAGCCGTGCGGTGTAATTGTTGCTCTCTCCGAAACTTGGCCAGAGTTCCA 108117
QY      172 GlyAlaLeuCyValLeuGlnGlnGlyGlyGlnThrGlnIleValLeuArgValGlyGln 191
      108118 GGTGTGCTGTGTGTGCTTTCAGAGAGCTGGGAGGACAACTGGTGTGACAACTGGAGAG 108177
QY      192 AlLeuLeuLysGlnValArgLeuProLysSerPheLysGlnAspValValLeuValMetAla 211
      108178 CACTTGCTTACGAGCCGTCCACTCAAGCGGTTCAGAGAGATGTGGCTTACGACATGGCA 108237
QY      212 LeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThr 231
      108238 TTGGCTTATGTGACATATCAAGGATGCTATGGCAGCAAGCCCTCCAGATGTATCTGC 108297
QY      232 GlyTyrGlnPheValGlnGluAlaLeuLysLeuGln----- 244
      108298 TGTGTGAGAGGTGCTTGAAGAGGCACTGAAGCTCCTCGAGGTGAACCTTGTCTTAAG 108357
QY      244 ----- 244
      108358 TTGCAGATCCAAATATGTCTTAAGGTGCTGTGTATCATATCAGTTGTTTAACTTAATGTG 108417
QY      244 ----- 244
      108418 ACATGTTCAGTTTAGGGCGAAATATGTGTGCACTACATATTTTGTAGATGAACA 108477
QY      244 ----- 244
      108478 TGCTCCAGTTTATATATATTGTTGTGTGTATACCATTAACACTTGTGTTTCTTGTGT 108537
QY      244 ----- 244
      108538 AAACAAATATCATTTGGGTAGACATTATGTTCAATAGTACTTTCATCTGCACTTAGATCT 108597
QY      244 ----- 244
      108598 CAGTAGCCCTTGATCATGATCATGTTATTAAGAACCTGAACCATTTGGCTACATCATGCG 108657
QY      244 ----- 244
      108658 ACAGCATTAATACTGTTTCACTGTAAACCAATCTGTATGAACCTATTAGAATGCT 108717
QY      244 ----- 244
      108718 GTATAGAGATATGAATATGCTCTAGTTTTCCTTTCTCTCTTATTTTAAATATATC 108777
QY      244 ----- 244
      108778 ATCCCATATTGCAAGCTGTACACAAAAAATGTCTGTGCAATTTTACGTGTGAGTT 108837
QY      244 ----- 244
      108838 ATTTTCACTGTTATTATGATCAGCGTGTGAAGTTTGTAGTCTGTCTTGTATGATGAGA 108897
QY      245 -----GlnGlnGlnValAspSerLe 251
      108898 TATCTTGGTAAAGCACTAATATGAAACATTTTACTTATAGAGATGTGTCAAGCAATCT 108957
QY      251 ValProAspLeuArgAlaGlnIleAspGlnThrLeuGlnGlnIleThrProArgTyrVal 271
      108958 TGCACCTGAACTGCTTGCACAGATTATGAAATCTTTGGAGAGATATACCTGCTGTCTGT 109017
QY      271 IleGlnLeuLeuGlyLeuProLeuGlnLysAspTyrAlaIleLysArgLeuAsnGlyLe 291
      109018 ATTGAGCTTCTTGTCTCTCTACTGATGAATAAAACATTAATAACGCCAAGAGGTCT 109077

```

QY	291	uSerGIyValaAgaenIIeLeuTrpSerValGIyGIyGIyValASerAlaLeuValGI	311
Db	109078	GCAGAGTGCAAAAACATATTTGGAGTGTGGCAGGGGGGTATTTGCTAGTTGGAGG	109137
QY	311	YGIyLeuTrpArGIyLysPheMetAenGIuAlaPheLeuArgMetThrAlaGIuGI	331
Db	109138	AGCATTTTCTCGTAGGCCCTTCATGAATAGAGCTTTCTTCGAGATGACATCAGCTAGCA	109197
QY	331	nVal	332
Db	109198	GGTTTGTGCTACCAAGTTTAATGAGCAACATCTAGTGTAGTGTGCTCATTTTAC	109257
QY	333	AspLeuPh	335
Db	109258	TTGAACACAGTTTTAGAACTTTCACATCTTCTGTTATATCTCAGATGAGATTTCTT	109317
QY	335	eValAlaThrProSerAenIIeProAlaGluSerPheGIuValTYrGIuValAlaLeuAl	355
Db	109318	CTCTTAACACCGAATATGACATACACCTGATGTGTTAGATTTATATAGTGTCCACTTGC	109377
QY	355	AlaLeuValAlaGIuAlaPheIIeGIyLysPProIIaLeuLeuGIuAlaAspLysGI	375
Db	109378	CCACATTTGCTCAAGCAATTCAGATAAGAAAGGCCAACATTCATCATATGATGSCAGATGAAC	109437
QY	375	nPheGIuGIuLeuGIuGIuAlaLysValMetAlaMetGIuIIeProAlaMetLeuTYrAs	395
Db	109438	CTTCGAAACAACCTTCAGAGATTCATATAGTTCCTCAA	109479
QY	395	pThrArgAenAntTrpGIuIIeAspPheGIuLeuGIuArgGIyLeuCyAlaLeuLeuII	415
Db	109480	TTATGAAAC-----GAGATGACCTCCGATTAGAAAGGCCACTGCTGCTACTTGT	109533
QY	415	eGIyLysValaAspGIuCyArgMetTrpLeuGIyLeuAspSerGIuAspSerGIuTYrAr	435
Db	109534	GGGGCATATTTCGAATTCGAGAAATGTGCTGGAATTTGATTAATGAGATCCCTCGCATATAG	109593
QY	435	GAnProAlaIIeValGIuPheValLeuGIuAenSerAen--ArgAspAspAenAspAs	454
Db	109594	GGACCCCAAAATATATGAGATTGTGTGTAATACTTAGCATGAAAGAGAGATATCT	109653
QY	454	pLeuProGIyLeuCyAlaLeuLeuGIuThrTrpLeuAlaGIyValValPheProArgPh	474
Db	109654	TTTTCCAGGCGTGTGCAGCTCTTGAGACATGCTGTCTCCGAGGTATTTCCAAAGGAG	109713
QY	474	eArgAspThrLysAspLysPheLysPheLysGIyAspTYrTYrAspAspProMetValIIe	494
Db	109714	CGAGATATCCCGAGGAGCAGCATTTGACATCGGAGACTATTACGACAGCCCAAAAGTTCT	109773
QY	494	uSerTYrLeuGIuArgValGIuValValGIuGIySerProLeuAlaAlaAlaIaThrMe	514
Db	109774	AGCTACTTAAGAAAGATGGAAGCGGTGTGCTCCCATTTTGTGCGAGCTCCGCTAT	109833
QY	514	AlaAspGIeGIyAlaGIu-----HisValLysAlaSerAlaMetGI	528
Db	109834	AGCAAAACTGTGCTCAAGCTACTGCTGCTGTGTAAGTAAATCAAGTCTCTTCA	109893
QY	528	nAlaLeuGIuLysValPheProSerArgTYrThrAspArgAenSerAlaGIuProLysAs	548
Db	109894	AGCTTTCCAGTAGGTTTTCCA-----TTGACAGAAACCTTGACATGTCCAGGAGGA	109947
QY	548	pValGIuGIuThrValPheSerValaAspProValGIyAsnAsnValGIyArgAspGIyGI	568
Db	109948	TATC-----	109951
QY	568	uProGIyValPheIIeAlaGIuAlaValArgProSerGIuAsnPheGIuThrAenAspTY	588
Db	109952	-----CCAAGTGAATGATCTGATTAATCT-----	109975
QY	588	AlaIIeArgAlaGIyValSerGIuSerValaAspGIuThrThyValGIuMetSerVal	608
Db	109976	-----CTTGA AAACTGGCCCAAGAAATGTGTCTGAGAGCGTATCATGATTTACAG	110028

QY	608	lAlaaspmetleuysgVlAlaseVallylglleuAlaAlaglyValAlaIlelyle	628
Db	110029	AAATGACCCCTTG-----AAGATATCTGCTGGTACACTGCTTCACCT	110073
QY	628	uIleSerleupheserGlnlyS-----	635
Db	110074	ATTTCAGGTGTTAGGTCTTCAGGTGCTTGCTGCTGTAAAGTCACTTCGCTCTTAAGGG	110133
QY	635	-----	635
Db	110134	CGAGTATGGGTCTGTGGCGGTGTGTGACTCATGTATGTGTAGTAAACCTGTATCAGATA	110193
QY	636	----TyrheleuysSerSerSerSerPheGlnArglySaPmetValSerSerMetG	654
Db	110194	GCATTACTTTCCTCCAAAGTCACTG-GGTTTT-----GCTTGATGCTCTTGACA	110240
QY	654	userAspValAlaThrIle-----	660
Db	110241	AAAGCGCGTGAAGAAATATTCACATAAATGTGTGCCAGGCGCATAACTCTTATTCT	110300
QY	660	-----	660
Db	110301	ATATGTATGCACTCATGTAGAGTGACACATAAATGCTCATAGAACCGTATTATTG	110360
QY	661	-----GlySerVa	663
Db	110361	TTAAACGTGCGCTGTGGCCAGGGAATTAATCTTCACCTTTGTGTTGTGAAGGTCACG	110420
QY	663	lArgAlaSaPAsPserGlnAlaLeuProArgMetSaPAlaArgThrAlaGluAsnIleVa	683
Db	110421	GCGAGATGAAGGCCACTGGAAATCTCTTAGATGACGCGGAAGTTGGCTGAAGACATTGT	110480
QY	683	lSerlySTrGlnlySilelySerleuAlaPheGlyProAspHis-----	698
Db	110481	TGCGAGGTGGCAAGACATCAAGTCCAAAGGCTTTGGGGCGAGAAACACATGTCAAGCGCAT	110540
QY	698	-----	698
Db	110541	GCAGAGGTGAGGTGCCAAAAATTTCCATCTTATAGCTAGCGCTTTGTTGTTGTTG	110600
QY	699	-----ArgIleGlu-----	701
Db	110601	TCCCAACATATCCCAACACACACATGTTGCTGAGAGTTGAGAGGCACTTTGGCTGTGC	110666
QY	701	-----	701
Db	110661	GGTCTTTGAAAACTCGACCTGCGGGGGTAAAGACAGCCCCGAGCATTTGTATTAAGAAAA	110720
QY	701	-----	701
Db	110721	GACCTTTCACACAGGTGAGAAAAACCCCGAAACCTTGCCCCACCCATACACAGCGGCA	110780
QY	701	-----	701
Db	110781	TGGAAGCCCATGTGAAGACACACGACCGGGGCTGAGCCTTAGACCTTGCTTTGGCGT	110840
QY	701	-----	701
Db	110841	GGGACAGACGAGAAGATTTTTTTTAAACACAGGCTGAATTTGCTGCCACGGGGAGTCGAA	110900
QY	701	-----	701
Db	110901	CCGAGACCTGAGAGTGTACTCAAAACCACTAACCACTGACGCGCAATGTGGGGA	110966
QY	701	-----	701
Db	110961	CTTGCTCTTAATGTATGAATTAAGAAAAAGCAACAAAAATTTGTTAAACGTTAAAG	111020
QY	701	-----	701
Db	111021	TGCTTCGTCCTTCGAAGCATTAATTTCTGTGGGATTAATGACTTTCAGACGAAGGTTAT	111080
QY	701	-----	701

[illegible]

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
CR954204_1		100001			
CR954204_2		200001			
CR954204_3		300001			
CR954204_4		400001			
CR954204_5		500001			
CR954204_6		600001			
CR954204_7		700001			
CR954204_8		800001			
CR954204					
Ostreococcus tauri	Ostreococcus tauri strain OTH0595, *** SEQUENCING IN PROGRESS ***	884654 bp	DNA	linear	HTG 05-JUL-2005
CR954204.2	GI:70062289				
HTG: HTGS_PHASE2	green alga, whole genome shotgun.				
Ostreococcus tauri					
Ostreococcus tauri					
Eukaryote: Viridiplantae; Chlorophyta; Prasinophyceae; Mamiellales;					
Mamiellaceae; Ostreococcus.					
1 (bases 1 to 884654)					
Detelle,E., Ferraz,C., Rombaouts,S., Rouze,P., Worden,A.Z.,					
Patensky,F., Degroove,S., Echeynie,S., Cooke,R., Robbins,S.,					
Muys,J., Saey,Y., Jabbari,K., Bowler,C., Ball,S., Ral,J.P.,					
Bouget,J.Y., Piganeau,G., De Baets,B., Picard,A., Delseny,M.,					
Demalle,J., Van de Peer,Y. and Moreau H.					
Genome analysis of the smallest free-living eukaryote Ostreococcus					
tauri unveils unique genome heterogeneity					
Nature (2005) In press					
2 (bases 1 to 884654)					
Rombaouts,S., Detelle,E., Ferraz,C., Van de Peer,Y. and Moreau,H.					
Direct Submission					
Submitted (30-APR-2005)					
On Jul 7, 2005 this sequence version replaced gi:62990096.					

```

BP 44, Avenue Fontaune, 66651 Banyuls sur mer, France Tel:
(33) 468887309
Fax: (33) 468887398
Email: h.moreau@obs-banyuls.fr
----- sequencing/assembly
Institut de Genetique Humaine, UPR CNRS 1142
34396 Montpellier Cedex 5, France
----- annotation
Department of Plant Systems Biology,
Flanders Interuniversity Institute for Biotechnology (VIB), Ghent
University, B-9052 Ghent, Belgium
Web site: http://bioinformatics.pdb.ugent.be/genomes.php
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 884654: contig of 884654 bp in length.
Location/Qualifiers
1..884654
/organism="Ostreococcus tauri"
/mol_type="genomic DNA"
/strain="OTH0595"
/db_xref="taxon:70448"
/clone_lib="Ostreococcus tauri genomic shotgun library"
/note="assembly number 050606"

```

RESULT	11
CR954204_0/c	
WPCOMMENT	
Sequence split into 9 fragments	LOCUS CR954204 Accession CR954204
Fragment Name	Begin End
CR954204_0	1 110000

Alignment Scores:	
Pred. No.:	2.99e-33
Score:	702.00
Percent Similarity:	45.5%
Best Local Similarity:	84
Query Match:	17.3%
Length:	1100000
Matches:	220
Conservative:	116
Mismatches:	291
Indels:	112

DB: 14 gaps: 25

US-10-600-070B-2 (1-801) x CR954204_0 (1-110000)

OY 78 AspargProGluArgHisValProIleProIleAspPheTyrGlnValLeuGlyValIleGln 97

Db 25764 GACCCCGGACGAGGCGACAGGTGGCGCGCTCCGCTTCGTAATCAATCAATTTGGGATTTGGC 25705

OY 98 ThrHisPheLeuThr--AspGlyIleArgArgAlaPheGlnAlaArgValSerIysPro 116

Db 25704 CCGCGCGGAACGACGCGCGAAGCGCGCTCCACGAGCGCGCTCCCGCTGATGACGCGCCG 25645

OY 117 ProGlnPheGlyPheSerAspAspAlaLeuIleSerArgGlnIleLeuGlnAlaIle 136

Db 25644 CTTCCCGAGGGTTATCTCCGTGCACCTGCTGAAACACGATGATCTTACACGAGGCC 25585

OY 137 CysGlnIleHisLeuSerAspProArgSerArgArgGlyTyrAsnGlnIleLeuLeuAspAsp 156

Db 25584 GTGGCGGTGCTCAAGATGAAAGAGCGCTCGCCGACAGACGATGAG-----GAC 25537

OY 157 GlnGlnAlaIleThrValIleIleHisPheValProTyrAspValProGlyAlaLeuCysVal 176

Db 25536 ATTGCGCAAGGTGTGCTCACCCTCGGTGAAGCCCATCGCGCCCGCCCTCTGTGCTTA 25477

OY 177 LeuGlnIleGlnGlyGlyIleThrGlnIleValLeuLeuArgValGlyGlnAlaLeuLeuGln 196

Db 25476 CTTGAGAGCGCGCGCGAGTACGAATCCGTATAGGATTTGAGCCCGCTCGCGCAGTCC 25417

OY 197 ArgLeuProIleSerPheIysGlnAspValIleLeuValMetAlaLeuAlaPheLeuAsp 216

Db 25416 GTAGCCGAGCCAGCAGCACCAGAGAGTGTGCGCTACCGCTCGCTCGGTGTGAG 25357

OY 217 ValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGlnPheVal 236

Db 25356 TTCGAGCACAATGCGCTCATGCGCAATCCGCTCGCATCCGAGGGTTGCCAATGCTCC 25297

OY 237 GlnGlnAlaLeuIleLeuLeuGlnGlnGlnGlnGlyAlaSerSerLeuAlaProAspLeuArg 256

Db 25296 GAGATGGGATCGAAGACGCTCTCGAGTGTGGC--CGGTGTTCGCGCGGAAGTCTCT 25240

OY 257 AlaGlnIleAspGlnIleThrLeuGlnIleIleThrProArgTyrValLeuGlnLeuLeuGly 276

Db 25239 CGCAACATCTCGCGCTTATATACAGACGTGCGCTCCGGCGTATGTCTCGAATTTGTTGCG 25180

OY 277 LeuProLeuGlyAspAspTyrAlaAlaIleArgLeuAsnGlyLeuSerGlyValArgAsn 296

Db 25179 ATGCGCGCTC-----GAGCGCAAGTCGAGCGCGCCCTGGGCGCGCCCTTCGCACG 25126

OY 297 IleLeuIlePheSerValIleGlyGlyGlyAlaSerAlaLeuValIleGlyLeuThrArgGln 316

Db 25125 CTCATTGTGGACMAAGACCCGTCGACACGCTCGAA-----CAGCGCGCG 25081

OY 317 LysPheMetAsnGlnAlaPheLeuArgMetThrAlaIleGlnIleValAspLeuPheVal 336

Db 25080 GCGTTTCATGAAACAGACGAGCAATTACTACCGCGCGAAGAGCAGCTCGCATCTTCGTC 25021

OY 337 AlaIleThrProSerAsnIleProAlaGlnIleSerPheGlnValTyrGlnValAlaLeuAlaLeu 356

Db 25020 GACGCTCGGATTAATATCCCGCTGACTCGAGCAATGTATCAAAAGACGCGCTCGCGCAC 24961

OY 357 ValIleAlaGlnAlaPheIleGlyLysLysPheProHisLeuLeuGlnIleAspLysGlnPhe 376

Db 24960 GTCTGCGCGGGTGTCAATCGACCGAACCATGCTCATCTCGACGCGGACGCAAAATCTTG 24901

OY 377 GlnGlnIleGlnGlnIleAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspThr 396

Db 24900 GAGCAAAATTCAGCTCGCTCGACGCTCTCGGCGAATCAAGCCGCTTC----- 24853

OY 397 ArgAsnAsnIleTyrGlnIleAspPheGlyLeuGlnIleArgGlyLeuCysAlaLeuLeuIleGly 416

Db 24852 -----TCCGAGTGTGCGCGTGAAGCGCGCGGTGTGTCAAAATTTGCTTGGT 24808

OY 417 LysValAspGlnCysArgMetIlePheGlnIleLeuAspSerGlnAspSerGlnTyrArgAsn 436

Db	24807	CGAGTTGAGGAAAGCCGAGTACACTCTGGGTTTACAGCAAGACCGCT-----GAT	24757
Oy	437	ProAlaIleValGluPheValLeuGluIleuSerAsnAlaGlyAspAsnAspLeu-----	455
Db	24756	CCGAGCTCTGTGCAATACATC-----GAGGATCGCTCGCCGAATGGAGATCTCCGC	24706
Oy	456	ProGlyLeuGlySerLeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArg	475
Db	24705	GAGGGTATGTGTCCGATGGCGGATCATGTGGCTGATCGACTGCGCTTCCCGCTTCCGC	24646
Oy	476	AspThrLysAspLysAspLysPhe---LysLeuGlyAspTyrTyrAspAspProMetValLeu	494
Db	24645	GGGACAGAGCAACAACCCAGCCGACCATCGAGATTGGTTCAACACCCGCAAGCTTCAA	24586
Oy	495	SerTyrLeuGluIleValGluValValGlnGlySerProLeuAlaAlaAlaIleThrMet	514
Db	24585	GGATTGTGTGAGCGGATTTGAT-----	24565
Oy	515	AlaArgIleGlyIleGluIleHisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPhe	534
Db	24564	-----TCGATGCCCGCCGCTGTTCA-----	24544
Oy	535	ProSerArgTyrThrAspArgAsnSerAlaGluProLysAspValGlnGluThrValPhe	554
Db	24543	-----ATCCAAGCGCGCATCGAA	24526
Oy	555	SerValAsp---ProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIle	573
Db	24525	TCGGCGAAGCAGACCGCTGGCTTCCGGCGTCGGTTGG-----ATT	24487
Oy	574	AlaGluAlaValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGly	593
Db	24486	GCGACGCTTTTCTTCCGACTCGCGG-----GCGGTCAAGTTTGGT	24445
Oy	594	ValSerGluSerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLys	613
Db	24444	ATCTCCGAT-----ACCCAACGAGAGCGGTGCT-----	24415
Oy	614	GluAlaSerValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSer-----	630
Db	24414	GTCGCAAAAGTATGTGTCTTC---GCCGGTGAAGCTTGTCTCAACACGAATGAGCC	24358
Oy	631	---LeuPheSerGlnLysTyrPheLeuLysSerSerSerPheGlnArgLysAspMet	649
Db	24357	GGTGTGGCTCTCGCGAGGATACCGCGCGAGCCGCGCGGTGCTCCGTTGCACAGC	24298
Oy	650	ValSerSerMetGluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGlu	669
Db	24297	GCTCGCGGCGCGCTGTCCAACGTCAAGTTGTGAACGCGACGCAACCAACGCGCAGC	24238
Oy	670	AlaLeuPro-----ArgMetAspAlaArgThrAlaGluAsnIleValSerLysTrpGln	687
Db	24237	CCGGCGCCCGTGAACCACTGACGAAGGGTACCGCGAGAGATGATGTGGCAATGGCAG	24178
Oy	688	LysIleLysSerLeuAlaPheGlyProAspHisArgGlyLeuMetLeuProGluValLeu	707
Db	24177	ACGGCCAAAGCGGAGGCTTTGGGGGACAGCGCAACATCTGCCAGCTGGAGAAATCTTA	24118
Oy	708	AspGlyArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeu	727
Db	24117	GAGGGGCGCGATGTTGCACGATGTGGCAAAAGCGCGCGAGAGAGTCAAGCGCAGGATGG	24058
Oy	728	ValTyrAspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAspGly	747
Db	24057	GCGGAGGAGTACCAACTGAATGAATGATATCGACAC-----GTTCAACCCGTTTGA	24004
Oy	748	Thr---ArgAlaLeuValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuVal	766
Db	24003	ACTGAAGAAATTTTGTGTGAAGCAGACGCTCACTGAAGTGTGCGGTGTGAAGAT-----	23950
Oy	767	HisProLysAsnAsnAlaThrAspVal-----ArgThrTyrThrArgTyrGluVal	784

Db 102013 GATCATCAGGCTTAAACATTGATGACTTCCTCCGATTCATCCAGCAATTCAGCCACAC 101954
 QY 326 MeethAlaAlaGlnGlnValaAspLeuPheValAlaThrProSerAsnIleProAlaGlu 345
 Db 101953 TTAAACGCGGCTGAAACATTAAGTTGATGCT--GAAAGCAAGCCCTTG-- 101900
 QY 346 SerPheGluValaIleValaAlaLeuAlaValaAlaGlnAlaPheIleGlyLeuValys 365
 Db 101899 GCTGTGGCTACATACCTTAGCTGTTTATGCTTCATGCGCAGAGATTCACCCAGCCAG 101840
 QY 366 ProHleuLeuGlnAspAlaAspIleGlnPheGlnGlnLeuGlnAlaValysValMet 385
 Db 101839 CCCGCTTATTCATGTCAGCAAGCAAAATTCGATCGGTTGCTAAGCGCGCA-- 101786
 QY 386 AlaMetGluIleProAlaMetLeuIleYrAspThrArgAsnAsnTrpGluIleAspPheGly 405
 Db 101785 -----GATGTGCAT 101777
 QY 406 LeuGluArgGlyLeuGlnCysAlaLeuLeuIleGlyLeuValaAspGluCysArgMetTrpLeu 425
 Db 101776 TTAGACGACCTCTGTGTGGCTATTACTAGGCGAACTGAAGAACCCAGCGAGTTTA 101717
 QY 426 GlyLeuAspSerGluAspSerGlnIleYrAspAsnProAlaIleValaGluPheValaLeuGlu 445
 Db 101716 GAACTGAGCCAGCA--TACGAAGCTTTAGCCTTAAT--CGA 101678
 QY 446 AsnSerAsnArgAspAspAsnAspAspLeuProGlyLeuGlnCysIleLeuGlnIleThrTrp 465
 Db 101677 GAAAAATCTCAAGATTCACCCGATTTACGCGAGTTTGCTTATATGCGCAACATGG 101618
 QY 466 LeuAlaGlyValaValaPheProArgPheArgAspThrIleYrAspIleValysPheValysLeuGly 485
 Db 101617 CTGCAGAAATGAAGATTTTCCCTCCATTTCCGCGATTGTGCCAGACAGCAAGCTTCCCTGAAA 101558
 QY 486 AspTyrIleAspAspProMetValaLeuSerTyrLeuGlnArgVala----- 500
 Db 101557 GATTACTTGTCTAATCAACAAGATACAAAGCGATTTTGAAGCCTTGCCCAACGACGCGGAA 101498
 QY 501 -----GluValaGlnGlySerPro 507
 Db 101497 ACCCATTAATGAATGGCTGTAAATTAACCCGCAATCGTTTCTCAACCCAGGCGCAATCT 101438
 QY 508 LeuAlaAlaAlaIleThrMetAlaArg-----IleGlyValaGluHis-----ValIys 523
 Db 101437 TACTCTGGAGGAACGCCAGTCCCAACGTCCTCCGTAAGGAGCGAACAGCCAGAGAA 101378
 QY 524 AlaSerAlaMetGlnAlaLeuGlnIleValaPheProSer-----ArgTyrThrAsp 540
 Db 101377 GCGTCCACAGAACAGATTCGCCCAACGTAATCATCCATCAGAAATTCGCGAGTTTCAT 101318
 QY 541 ArgAsnSerAlaGluProIleValaAspValaGlnGluThrValaPheSerValaAspProValaGly 560
 Db 101317 CAAAAACGAACCCCTGATCCCGAATTAACAGAAACATCAACACAGAGACAGAGTCT 101258
 QY 561 AsnAsnValaGlyArgAspGlyIle----- 568
 Db 101257 TCAAAATTTTAACTGCTAGAGAAATATATTCAGACACAGATGCTTACCTGACCAATTAT 101198
 QY 569 ---ProGlyValaPheIleAlaGluAlaValaArgProSerGluAsnPheGluThrAsnAsp 587
 Db 101197 CCACACAGATCCCTGTAGAACCGCGCAGAGACT----- 101162
 QY 588 TyrAlaIleArgAlaGlyIleSer-----GluSerSerValaAspGluThrThrVal 604
 Db 101161 -----GTTACGCCGGGGTAAAGTATTATACCAATCGACCCCTCCACGCAAACTCCT 101108
 QY 604 ----- 604
 Db 101107 AAACGAGAGACGCAAGAAACCAAGCAGGCTGTTCACACAGAGACACAGATTCATCAG 101048
 QY 605 -----GluMetSerValaIleAspMetLeuIleGluIleValaValysIleLeuAlaAla 622
 Db 101047 CAACGCCAACCTCCTACGACATCTTAGCGCGGAAACAGATTAATCTTGGATATGTTTG 100988

QY 623 GlyValaAlaIleGlyLeuIleSerLeuPhe-----SerGlnIleYr--Phe 637
 Db 100987 GGT---TCTTGGGCGGATATATTATGTTCTGCTGATAGTCTCAACGACTTTGGGTGG 100931
 QY 638 LeuIleYrSer-----SerSerPheGlnAlaGlyIleAspMetValSerMet 653
 Db 100930 TTAAAGAAATGATATCTTCCACACCATCTTTTACAGAGTGACCAATTTATCGATTAGATT 100871
 QY 654 GluSerAspValaAlaThrIleGlySerValaArgAlaAspAspSerGluAlaLeuProArg 673
 Db 100870 ACTCAACACCTTTAGAGATTCCTGACAAATAATGCCAGATACATCCCCAGAGTGACT 100811
 QY 674 MetAspAlaArgThrAlaGluAsnIleValaSerIleTrpGlnIleValysIleSerLeuAla 693
 Db 100810 CTCACAGAAAGAACCGCAAGAAATTAATGAAAATTTGGTGGCTACCAACGAGTGCT 100751
 QY 694 PheGlyProAspPheArgIleGluMetLeuProGlyValaLeuAspGlyArgMetLeuValys 713
 Db 100750 TTAGCGGCTGAACATTAATTAAGAGTTTAAACGAGATTTTAACTGTTACAGCGTTATCT 100691
 QY 714 IleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValaIleYrAspTyrThrLeu 733
 Db 100690 CAATGCGGCTAATTCCTTGCAAGATTAAGACAGACATCGTCATCGAGATTCAGT--- 100634
 QY 734 LeuIleYrLeuSerValaAspSerValaThrValSer---AlaAspGlyThrArgAlaLeuVal 752
 Db 100633 CATAGTGTCAAGGTAGATTCATCCATGATGATGATGATGATGATGATGATGATGATGATG 100574
 QY 753 GluAlaThrLeuGlnGluSerAlaCysIleSerAspLeuValHis-----ProGluAsn 770
 Db 100573 GGGGCTACAGTCAAGAG--TTAACCCAAATTTATGAGAAATGGGCAAAAGGG 100523
 QY 771 AsnAlaThrAspValaArgThrIleYrThrIleArgTyrGluValaPheTrpSerIleYrSerGly 790
 Db 100522 AAGTCTTCTGACGAAAGA--TTACGTGTACGCTATGAAATGATTCGACAAAGTGTATT 100466
 QY 791 TrpIleVal 793
 Db 100465 TGGCGGATT 100457

RESULT 13
 BA000019_33/c
 WPCOMMENT
 Sequence split into 65 fragments LOCUS BA000019 Accession BA000019
 Fragment Name Begin End
 BA000019_00 1 110000
 BA000019_01 100001 210000
 BA000019_02 200001 310000
 BA000019_03 300001 410000
 BA000019_04 400001 510000
 BA000019_05 500001 610000
 BA000019_06 600001 710000
 BA000019_07 700001 810000
 BA000019_08 800001 910000
 BA000019_09 900001 1010000
 BA000019_10 1000001 1110000
 BA000019_11 1100001 1210000
 BA000019_12 1200001 1310000
 BA000019_13 1300001 1410000
 BA000019_14 1400001 1510000
 BA000019_15 1500001 1610000
 BA000019_16 1600001 1710000
 BA000019_17 1700001 1810000
 BA000019_18 1800001 1910000
 BA000019_19 1900001 2010000
 BA000019_20 2000001 2110000
 BA000019_21 2100001 2210000
 BA000019_22 2200001 2310000
 BA000019_23 2300001 2410000
 BA000019_24 2400001 2510000
 BA000019_25 2500001 2610000
 BA000019_26 2600001 2710000


```

QY 501 -----GluValValGlnGlySerPro 507
DB 1497 ACCACTAATGAATGGGTGTAATTAACGCCCATCGTTTCTGACCCAGGGCAATTCT 1438
QY 508 LeuAlaAlaAlaThrMetAlaArg-----IleGlyValGlnHis-----ValLys 523
DB 1437 TACTCTGAGGAACGCCAGTCGCGCAAACTGCCGTAGGAGGCAACAGCCAGGAGAA 1378
QY 524 AlaSerAlaMetGlnAlaLeuGlnLysValPheProSer-----ArgTyrThrAsp 540
DB 1377 GCGTCCCAAGAACAGCTGCCCAACAGTCATCATCAGACAGTAATTAATCGCGATTTCAT 1318
QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValaAspProValGly 560
DB 1317 CAATAACAGAACCCCTGATCCCGAATTACGAAACATCAAAACACAGAAACACAGAGTCT 1258
QY 561 AsnAsnValGlyArgAspGlyLys-----568
DB 1257 TCAAAATTTTCAACTGCTAGAGAAATATATCGAACACAGATGCTTACACTGACAAATTAT 1198
QY 569 ---ProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
DB 1197 CCACCAAGATCCCTGTAGAACCGCCAGCAGACCT-----1162
QY 588 TyrAlaIleArgAlaGlyValSer-----GluSerSerValaAspGluThrThrVal 604
DB 1161 -----GTTACGCCGGGGGTAGTGTTATACCAATCGACCCCTCCACGGCAAACTCCT 1108
QY 604 -----604
DB 1107 AAACGAGAGACGCAAGAACCAAGCAGATTGTCAACAGAGACACAGTATTTCATCAG 1048
QY 605 -----GluMetSerValaIaAspMetLeuLysGluAlaSerValLysIleLeuAlaAla 622
DB 1047 CAACGCCAACCCCTCACCCTGACCTAGCGCCCAAAACAAGTATCTTGATGATTTTG 988
QY 623 GlyValAlaIleGlyLeuIleSerLeuPhe-----SerGlnLysTyr---Phe 637
DB 987 GGT---TCTTGGGTGGGATTAATTATGTTCTGGCTGATCTCAACGACTTTGGTGT 931
QY 638 LeuLysSer-----SerSerSerPheGlnArgLysAspMetValSerSerMet 653
DB 930 TTAAGAATGTATTTCTCCAGACACCATCTTCAACAGTGACGAATATGATTCAGATT 871
QY 654 GluSerAspValaIaThrIleGlySerValaArgAlaAspAspSerGluAlaLeuProArg 673
DB 870 AGTCAACACCTTAGAGATTCTGACAAAATGCCCCAGATACATCCCAAGAGTGAGT 811
QY 674 MetAspAlaArgThrAlaGluAsnIleValSerLysTyrGlnLysIleLysSerLeuAla 693
DB 810 CTCACAGAAACCGCAGAGAAATTAATGAAATGCTGGCTACCAACGATAGTCT 751
QY 694 PheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuLys 713
DB 750 TTAGCGCGCTGAACATAAATTGAGAGTTTAAACGAGATTTTAACTGGTTCAGCGTTATCT 691
QY 714 IleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeu 733
DB 690 CAATGGGGCTAATTCCTTGCCTTGCAGAGATMAAGACAGACATCGTCATCGAAGATACAGT-- 634
QY 734 LeuLysLeuSerValaAspSerValaIthValSer---AlaAspGlyThrArgAlaLeuVal 752
DB 633 CATAGCTCAAGTAGTACATCATCATCAATTCGACATTAATCCCAATCTGTGCAAGGTG 574
QY 753 GluAlaThrLeuGlnGluSerAlaCylLeuSerAspLeuValHis-----ProGluAsn 770
DB 573 GGGGCTACAGCTCAGAGAG-----TTAACCAATTTTATGATGAGATGGCAAAAGGG 523
QY 771 AsnAlaThrAspValaArgThrTyrThrThrArgTyrGluValPheTyrSerLysSerGly 790
DB 522 AAGCTTCTGACGCAAGAA---TTACGTAGCGTATGATTTGATTCGACAAAGATGATATT 466
QY 791 TrpLysIle 793

```

```

DB 465 TGGCGGATT 457
RESULT 14
BA0000045_36
WPCOMMENT
Sequence split into 47 fragments
Fragment Name Begin End
BA0000045_00 1 11000
BA0000045_01 100001 21000
BA0000045_02 200001 31000
BA0000045_03 300001 41000
BA0000045_04 400001 51000
BA0000045_05 500001 61000
BA0000045_06 600001 71000
BA0000045_07 700001 81000
BA0000045_08 800001 91000
BA0000045_09 900001 101000
BA0000045_10 1000001 111000
BA0000045_11 1100001 121000
BA0000045_12 1200001 131000
BA0000045_13 1300001 141000
BA0000045_14 1400001 151000
BA0000045_15 1500001 161000
BA0000045_16 1600001 171000
BA0000045_17 1700001 181000
BA0000045_18 1800001 191000
BA0000045_19 1900001 201000
BA0000045_20 2000001 211000
BA0000045_21 2100001 221000
BA0000045_22 2200001 231000
BA0000045_23 2300001 241000
BA0000045_24 2400001 251000
BA0000045_25 2500001 261000
BA0000045_26 2600001 271000
BA0000045_27 2700001 281000
BA0000045_28 2800001 291000
BA0000045_29 2900001 301000
BA0000045_30 3000001 311000
BA0000045_31 3100001 321000
BA0000045_32 3200001 331000
BA0000045_33 3300001 341000
BA0000045_34 3400001 351000
BA0000045_35 3500001 361000
BA0000045_36 3600001 371000
BA0000045_37 3700001 381000
BA0000045_38 3800001 391000
BA0000045_39 3900001 401000
BA0000045_40 4000001 411000
BA0000045_41 4100001 421000
BA0000045_42 4200001 431000
BA0000045_43 4300001 441000
BA0000045_44 4400001 451000
BA0000045_45 4500001 461000
BA0000045_46 4600001 4659019
Continuation (37 of 47) of BA0000045 from base 3600001 (BA0000045 Gloeobacter violaceus P
Alignment Scores:
Pred. No.: 3,85e-17 Length: 110000
Score: 449.00 Matches: 197
Percent Similarity: 40.4% Conservative: 114
Best Local Similarity: 25.6% Mismatches: 287
Query Match: 11.1% Indels: 172
DB: 1 Gaps: 31
US-10-600-070B-2 (1-801) x BA0000045_36 (1-110000)
QY 61 PheAlaThrAlaThrThrAlaThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
DB 89225 TTTTTCGGCGCCATGATGATTCGCTCATCCCTCCCTG-----89263
QY 81 GluArgHisValaProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100

```


Db 89264 -----CCGCTGAGCTACTACAGATCTCGGAGTGCCTCCCACTGTC 89305
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIysProProGlnPheGly 120
 Db 89306 ACTTACAGAGAGGAGGAGCTGCTTTGCGGARTGCTTGGCGCCAGCGCGAG 89365
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleGluAlaIleAlaCysGluThrLeu 140
 Db 89366 TTTTACGCGCGCGCTCCGCTCCGCGAGACCTGCTGCGCGAAGCTGCACAGCGCTG 89425
 QY 141 SerAsnProArgSerArgArgValGluTyrAsn--GluGly-----LeuLeuAsp 155
 Db 89426 GCGGATCTCCGCGAGCTGTATCATCGAGGGCAACAAGATTGCTTCTCGAC 89485
 QY 156 AspGluGluAlaThrValIleThrAspValProTyrAspIysValProGlyAlaLeuCys 175
 Db 89486 AGCAGCCACGCGCGGTG-----GGGCTGCTGTTT 89515
 QY 176 ValIleGluGlnIleGlyGluThrGluIleValIleuArgValGlyGluAlaLeuIys 195
 Db 89516 CTGTACGAAGTGGGCGCAATACCAAGCTCATCGAGCGCAAGGAGCTTACGCGTA 89575
 QY 196 GluArgLeuProIysSerPheIysGlnAspValValIleuValMetAlaLeuAlaPheLeu 215
 Db 89576 GCGGATCAGCCG-----GATACCGCGCTGGTCTCGGCTCGCCACAG 89620
 QY 216 AspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPhe 235
 Db 89621 GCCCTGGCCAGAGAGCTTACCCG-----CAGGGCAATCTCACCTGGCCACCTCGAA 89674
 QY 236 ValGluGluAlaLeuIysLeuLeuGlnGluGluIleValIleSerSerLeuAlaProAspLeu 255
 Db 89675 GTTGAAGAGCTTGTGAGATTGTCG-----GGCGCGATTGCTCGAAGCCG--GTC 89725
 QY 256 ArgAlaGlnIleAspGluThrIleGluGluIleThrProArgTyrValIleGluLeuLeu 275
 Db 89726 CACGAGAGTTGAGACTGCTCTCAAGCGCTGGCGGCCGAGGAGATTGTCAGCTGTG 89785
 QY 276 GlyLeuProLeuGluAspAspTyrAlaAlaIleArgLeuAsnGlyLeuSerGlyValArg 295
 Db 89786 GCG--GAGCGGCGCAGCCCGCTCCCGCCAGCGCGAGAGAGCATGGCTCGCG 89842
 QY 296 AsnIleLeu-----TrpSerValGlyGlyGlyValIleSerAlaLeuValGlyGly 312
 Db 89843 GCCCTGCTTCGCGAGCGCGAGCGCATCGAAGCGGACGCAACATCA-----TCGGCG 89896
 QY 313 LeuThrArgGlu-----LysPheMetAsnGluAlaPheLeuArgMetThrAla 329
 Db 89897 CTCAGCGCGAGAAATTGTTCAGATTGTCAGATCTGCGCTGCGCGTGCAGCGTGGC 89956
 QY 330 GluGlnValAspLeuPhe-----ValAlaThrProSerAsnIleProAlaGluSer 346
 Db 89957 GAGCACAAGAAATTGTTTCAGCGCGAAGCGCGCGCGCTCG-----CCGCGCGCCAG 90010
 QY 347 PheGluValIleThrIleValAlaLeuAlaLeuAlaGlnAlaPheIleGlyIysPro 366
 Db 90011 -----TACCTGCGCGCGCAGCGCAATGCGCTGCGGATTACTGAAGCTCGCC 90061
 QY 367 HisLeuLeuGlnAspAlaAspIysGlnPheGlnIleGluGlnIleAlaIysValMetAla 386
 Db 90062 CAGTGGTGGCGCGCGCGCGGAGCACTCATCAGCTCTTCAAGCGCCAG----- 90112
 QY 387 MetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeu 406
 Db 90113 -----GACGTCAATCTG 90124
 QY 407 GluArgGlyLeuCysAlaLeuLeuIleGlyIysValAspGluCysArgMetTrpLeuGly 426
 Db 90125 GAACCTGCGCTGTGGCTGCTGTGTGGCCAGGTGCAGAGCCCAAGAAACATCGAG 90184
 QY 427 LeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPheValIleGluAsn 446
 Db 90185 CGCTGCGCGAGAGAGCG-----GCGGTCAATTACATC---AAAAAC 90223

QY 447 SerAsnArgAspAspAsnAspAspLeuProGlyIleuCysIleLeuLeuGluIleThrTrpLeu 466
 Db 90224 CTGTCCAGATTCGCCCGGACCTGCTGCGGGCTGTGCGGTATACCACTGTGGCTG 90283
 QY 467 AlaGlyValAlaPheProArgPheArgAspThrIysAspIysIysPheIysLeuGlyAsp 486
 Db 90284 GCGGAGAGGTCTTTTCAGGTTTTTCGGACCGGCGCTCGGAGCACTATACCTGAGGCT 90343
 QY 487 TyrTyrAspAspProMetValLeuSerTyrLeuGluArgValGluValIleGlnIleSer 506
 Db 90344 TACTTTGCCCATCCAGAGTCCGCGGCTTCTCGAT-----GAC 90382
 QY 507 ProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGluIleValIysAlaSerAla 526
 Db 90383 CCGCAACCGCGCGGCTCTGCTCCCGGAACCGCGACCGCGCGCGATTGCGCGCG 90442
 QY 527 MetGlnAlaLeuGlnIysValPheProSerArgTyrThrAspArgAsnSerAlaGluPro 546
 Db 90443 GCGTCGTG-----GSCACCATCGAACC 90466
 QY 547 LysAspValGlnGluThrValPheSerValAspProValGlyAsnAsnValGlyArgAsp 566
 Db 90467 GCGCGCTG-----CTGACCGCGGTGGG----- 90490
 QY 567 GlyGluProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsn 586
 Db 90491 -----GCCAGGCGCGCGCGCGCGC----- 90511
 QY 587 AspTyrAlaIleArgAlaGlyValSerGluSerSerValAspGluThrThrValGluMet 606
 Db 90511 ----- 90511
 QY 607 SerValAlaAspMetLeuIysGlnAlaSerValIysIleLeuAlaIleAlaIle 626
 Db 90512 -----GCCAATCTGCTCGCCCGAGCGCTGCTGCGCGCTCGCTCGCGAATT 90565
 QY 627 GlyLeuIleSer-----LeuPheSerGlnIleTyrPheLeuIysSerSerSer 643
 Db 90566 TTCCTCTCGCGAGCGGCTGCTGCTACCCAGCGCGAGAGCGCGCGAGCTGACCG 90625
 QY 644 PheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIleGlySerVal 663
 Db 90626 GCGCGCAAGCC-----GTGCTCTCGGTGCAGGCTCCGCTCATTCGCGCGCC 90679
 QY 664 ArgAlaAspAspSerGluAlaLeuProArg----- 673
 Db 90680 CCGGCGCGCGCGAGCGCGCTGCCACCGAGCGGTGCGCGAGCAACAGCTTCC 90739
 QY 674 MetAspAlaArgThrAlaGluAsnIleValSerIleTyrGlnIleIleIysSerLeuAla 693
 Db 90740 ACCGAGCGCAATTCGCCGCG--ATGCTCAAAAGTGCAGAACGCGCAAGCGAGGCC 90796
 QY 694 PheGlyProAspHisArgIleGluMetLeuProGluValIleuAspGlyArgMetLeuIys 713
 Db 90797 CTGCTTCGCGAGCACCGCACCGCGCATGTCAGATGCTCACAGGCTCCCGCACGCG 90856
 QY 714 IleTyrThrAspArgAlaAlaGluThrAlaGlnIleGluIleuValIysAspTyrThrLeu 733
 Db 90857 GTCGCGCACAAAAAGTTCAGAGAGCGCGCGCGGAGATTCGAAATTCAAGCTTC 90916
 QY 734 LeuIysLeuSerValAspSerValThrValSerAlaAsp-----GlyThrArgAlaLeu 751
 Db 90917 AAGATCTCAAAATTCAGAGGTG-----GCCAGCGCGCGCGCGAGCGGTGGCG 90967
 QY 752 ValGluAlaThrIleGluGluSerAlaCysLeu-----SerAspLeuValHisProGlu 769
 Db 90968 GTGACGCGCGAGTCAACGAGTGCAGATCTCTACACGAGCAATAGCTGCGCGCTCG 91027
 QY 770 AsnAsnAlaThrAspValArgThrTyrThrArgTyrGluValPheThrSerIysSer 789
 Db 91028 CGCTCTACGAC-----GCGCCCTACCGGGTGTACTCTCCGTGTCAAGAGCCCTGCC 91081

Db	91082	G3CTGGCGATCGAGAGATGAAAGTGTG	91111
RESULT 15			
LOCUS	AF421196		
DEFINITION	AF421196	2469 bp	DNA
ACCESSION	Synechococcus sp. PCC 7942 cell division protein Ftn2 gene,	linear	BCT 10-SEP-2002
VERSION	complete cds.		
KEYWORDS	AF421196		
SOURCE	AF421196.1	GI:16226083	
ORGANISM	Synechococcus elongatus PCC 7942		
REFERENCE	Synechococcus elongatus PCC 7942		
AUTHORS	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.		
TITLE	1 (bases 1 to 2469)		
	Koksharova,O.A. and Wolk,C.P.		
	A Novel Gene That Bears a DnaJ Motif Influences Cyanobacterial Cell		
JOURNAL	Division		
PUBMED	J. Bacteriol. 184 (19), 5524-5528 (2002)		
REFERENCE	12218043		
AUTHORS	2 (bases 1 to 2469)		
TITLE	Koksharova,O.A. and Wolk,C.P.		
JOURNAL	Direct Submission		
	Submitted (18-SEP-2001) Plant Research Laboratory, Michigan State		
	University, DOE Plant Research Laboratory, Batc Lansing, MI 48824,		
	USA.		

CDS

ORIGIN

Alignment Scores:

Pred. No.:	5,4e-19	Length:	2465
Score:	446.00	Matches:	203
Percent Similarity:	38.8%	Conservative:	129
Best Local Similarity:	23.7%	Mismatches:	292
Query Match:	11.0%	Indels:	232

US-10-600-070B-2 (1-801) x AF421196 (1-2469)

Qy 4 Leuserhsva1g11ieg1yleuserpProphcInLeucyarguePropho1aThr 23
Db 161 CTCCTCAAGAAAGGGTTTCCTCGCGC--TATTCCTGCATCATCCCCCTGTCTGCT 21
Qy 24 Thyr1sLeutarg1aGserSeri1aSenthrSerThrThr1ecySera1aSer1yStrp1a 43
Db 218 ACCGAGATAC--GCCATCATCTACACGAA--TTGGTAAAGTGAAGGTCAAGTTGGAT 27
Qy 44 Aaphr1euleuserAepPheanPheThrSer1aSpSerSerSerSerPhealathr 63

Dd	274	GATGAAGTCGGCTCAGACTGGCCATCCGAGATCTGGTGGGTTCT-----	318
Oy	64	AlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgProGluArgHis	83
Dd	318	-----	318
Oy	84	ValProIleProIleAspPheTyrGlnValLeuGlnYlaIaGlnThrHisPheLeuThrAsp	103
Dd	319	GTGGGTATTCCTCTCCGATTAATCAACGAATTCCTGTGTGGAGCGTGAAGCTCGGCAGAC	378
Oy	104	GlyIleArgArgAlaPheGluAlaArgValSerIleAspProGlnIlePheGlyPheSerAsp	123
Dd	379	AAACTTCGCCAAGAACTACCGCGATCGGCTCAACAACATCGGCTCCCATGATGATTTCAGAG	438
Oy	124	AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaAcYsgIuThrLeuSerAsnPro	143
Dd	439	CTGCATTTGACGGCGCGCGCAACTCTCGAAGACAGCATTCGTGATCGTAGTAAGATCCC	498
Oy	144	ArgSerArgArgGluTyrArgGln-----GlyLeuLeuAspArgGlu	158
Dd	499	GAACAGCGCGATCGCTACGATCGCCGCTTTTTCAGGCGGCTGTGAAGCATTTGAACCA	558
Oy	159	AlaThrValIleThrAspValProTyrAspIleValProGlyAlaLeuCyValLeuGln	178
Dd	559	AGCCTAAGAACTCGAAGAC-----TGCACAGCAATT--GGAAGCTCGTGAATCTCTGTG	609
Oy	179	GluGlyGlyGlnThrGlnIleValLeuArgValGlyGluAlaLeu-----	194
Dd	610	GAATTTGGGGGAATACAGATCGGTTTCCGATCGGCTGAGAACTCTCGCAGACTACGAC	669
Oy	195	-----LysGluArgLeuProLysSerPheIleGlnAspAlaValLeuValMetAlaLeu	212
Dd	670	GCGAGCCGACAGATACCGCATCACTTCGCGCGGGGTATATTCGCTTGGCGATCGACTA	729
Oy	213	AlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGly	232
Dd	730	TCCACGACATCCCTCGGTCAAGATGCCCTCAG-----	762
Oy	233	TyrGluPheValGluGluAlaLeuLysLeuLeuGlnGluGluYala-----	248
Dd	763	-----CAGGGTCTGACGACGAGCGGCCCCAGCACTTGGC	798
Oy	249	-----SerSerLeuAla-----ProAspLeuArgAlaGlnIleAsp	260
Dd	799	CGCAGCCAGTCTGGCTTACGCCGATCATACGCGCTTCTTAACATGACTGAACTTCAC	858
Oy	261	GluThrLeuGlnGluIleThrProArgTyrValLeuGlnLeuLeuGlyLeuProLeuGly	280
Dd	859	CAGAAACAGAGACAGCTACGGCTTATTCGATTTTGAAGCGGTTGGCCAGCCCTTACT	918
Oy	281	AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeu-----	298
Dd	919	GCCGAT-----ACGATTCGCGACAGGAGTTGCTGTGTGACAGCGAATTTGGACGAC	972
Oy	299	---TrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLys	317
Dd	973	CGGACGGGCAATTGAAGGCCCTGGGAGATGATGCG-----TCGGGCGTGACCTTGATPAC	1026
Oy	318	PheMetAsnGluAlaPheLeuArg-----MetTrrAlaGlnGlnVal	332
Dd	1027	TTTTTG-----ATGTTTCTCCACAAATTCGCGCGATATCTGACCTCGCTGAACACAG	1080
Oy	333	AspLeuPhe-----ValAlaThrProSerAsnIleProAlaGlnSerPheGluVal	349
Dd	1081	TTTGCTTTTGAATCGAAGCGCGTGGGCGCTCG-----CCGGCTGGAGCTT-----	1128
Oy	350	TyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLysIleProHisLeuLeu	369
Dd	1129	---TTTGGCTGCTACACCTGATTTGGCGGGGGCTTTTGGATCAACCAACCTCGTTATC	1185
Oy	370	GlnAspAlaAspLysGlnPheGlnGlnLeuGlnIleLysValMetAlaMetGluIle	389
Dd	1186	CATCGCGCGACGTTGCTTGCATGAATCTCAAGAC-----	1221

THIS PAGE BLANK (USPTO)


```
/canelation="MEALSHVIGISLPOLCLRPATTKLRSHNTSTTICSAKMWAD
RLSDPNFSSSSSPATTTATVLSIPSIDREHPPIIDRYOVLGAQTHET
DTRBARPEARVSRPOGPDMDALISROIOAACTLSNPSRBRNGLDDECAT
VITDVPMKVPKALCYLOEGETIYLRVBEALIKERLPSPFQDVLVWALAFLVDS
RDMALDPDPIITGYEBERALILQEBGASLAPDLRAQIDTLEITPRVYLEILG
LPLGDDYAAKRLINGSLGVRNLIWSVGGASLAVGLTRKFNMEFLRMTAEQVUL
FVTPNIPAPESFEVYEVVALVAQAFIGKPHILOADKFOLOQAKMAEINIPAM
LVOTRNNWEIDFELBERGICALIIGVDECMWGLDSEDSQYENPAIIVERFVLENSRD
DNDLPGLCLEETMYLGVPRPRDCKKFKLGPYDDPMVLSTLEREVVQSGPL
AAATATRIAEHVRKASAMQALOKVFPSTRTDPSAPRHOVQETVSVDPVGNVRD
GEBGVFAEVRPSENFTNDVAIRAGVSSSVDETTEWMSVAMELKESVKLIAGV
ALGILIFSQKFLKSSSFORKDMWSMESDVATIGSVADQSEALPRMDATFANI
VSKWOKISLAFGPDHRIEMLPEVLDGRMLKWTDBRAAEATQOLGLVYDTLLSTVDS
VTVSADGTRALVETALIEBSACLSLDIVHPENNATDVRYTTRREVFWSKSGMKITEGSV
LAS"
misc_difference 474
/gene="ARC6"
/ntoe="results in G1u to G1u substitution"
/replace="a"
misc_difference 1831
/gene="ARC6"
/ntoe="results in G1u to Asp substitution"
/replace="t"
ORIGIN
Query Match 99.9%; Score 2402.8; DB 15; Length 2438;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAAGCTCTGAGTACGTGCGGATTTGATCTCTCCCATTCATATTAGCCGATTACA 60
DB 2 ATGGAAGCTCTGAGTACGTGCGGATTTGATCTCTCCCATTCATATTAGCCGATTACA 61
QY 61 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACCTCTTACATATCTGCTCGGCGAGC 120
DB 62 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACCTCTTACATATCTGCTCGGCGAGC 121
QY 121 AAATGGGCGGACGCTCTCTCTCGACCTTCAATTCACTCCGATTCCTCTCTCTCTCC 180
DB 122 AAATGGGCGGACGCTCTCTCTCGACCTTCAATTCACTCCGATTCCTCTCTCTCTCC 181
QY 181 TTGCGCACCGCCACCAACCGCACCTCTCTCTGCGCACCATATATGATGATGCC 240
DB 182 TTGCGCACCGCCACCAACCGCACCTCTCTCTGCGCACCATATATGATGATGCC 241
QY 241 GAAAGCCACGTCCTCCCATTCGATTTCTTACCAAGTATTAGAGCTCAAAACATTTCC 300
DB 242 GAAAGCCACGTCCTCCCATTCGATTTCTTACCAAGTATTAGAGCTCAAAACATTTCC 301
QY 301 TTAAACGATGGAATCAAGAGGATTTGAAGCTAGGGTTTGAACCGCGCAATTCGGT 360
DB 302 TTAAACGATGGAATCAAGAGGATTTGAAGCTAGGGTTTGAACCGCGCAATTCGGT 361
QY 361 TTGACGACGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG 420
DB 362 TTGACGACGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG 421
QY 421 TCTAATCTCGGCTTGAAGAGAGTACATGAAGTCTTCTTGAATGATGAAGAGCTTACA 480
DB 422 TCTAATCTCGGCTTGAAGAGAGTACATGAAGTCTTCTTGAATGATGAAGAGCTTACA 481
QY 481 GTCTATCTGATGTTCTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGCAAGAAAGT 540
DB 482 GTCTATCTGATGTTCTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGCAAGAAAGT 541
QY 541 GGTGAGACTGAGATGATTTCTTGGGTTGGTGAAGCTCTGCTTAAAGAGAGTTGCTTAAG 600
DB 542 GGTGAGACTGAGATGATTTCTTGGGTTGGTGAAGCTCTGCTTAAAGAGAGTTGCTTAAG 601
QY 601 TCGTTTAAAGCAAGATGTTGTTTAAAGTATGAGCGCTTGGCTTTCTGATGTTCTGAGGAT 660
DB 602 TCGTTTAAAGCAAGATGTTGTTTAAAGTATGAGCGCTTGGCTTTCTGATGTTCTGAGGAT 661
```

```
QY 661 GCTATGAGCATTTGATTCACCTGATTTTATTTACTGTTATGAGTTTGTGAGAAAGCTTTG 720
DB 662 GCTATGAGCATTTGATTCACCTGATTTTATTTACTGTTATGAGTTTGTGAGAAAGCTTTG 721
QY 721 AAGCTTTTACAGAGAGAGAGAGCAATAGCTTTGCAACCGGATTTAGTGACAAATTTGAT 780
DB 722 AAGCTTTTACAGAGAGAGAGAGCAATAGCTTTGCAACCGGATTTAGTGACAAATTTGAT 781
QY 781 GAGACTTTGAGAGAGATCACTCCGCTTATGTCCTTGAGACTACTTGGCTTACCGCTTGGT 840
DB 782 GAGACTTTGAGAGAGATCACTCCGCTTATGTCCTTGAGACTACTTGGCTTACCGCTTGGT 841
QY 841 GATGATTAACGCTGCGAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGCTC 900
DB 842 GATGATTAACGCTGCGAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGCTC 901
QY 901 GTTGGAGAGAGTGGACATCAAGCTCTGTTGGGGGGTTTGAACCGGTAGAAATTATGAAT 960
DB 902 GTTGGAGAGAGTGGACATCAAGCTCTGTTGGGGGGTTTGAACCGGTAGAAATTATGAAT 961
QY 961 GAGGCGTTTTTTCAGAAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCGAGC 1020
DB 962 GAGGCGTTTTTTCAGAAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCGAGC 1021
QY 1021 AATATTTCCAGACAGATCATTTGAAGTTTACGAAGTTGACATTCCTTGTGGCTCAAGCT 1080
DB 1022 AATATTTCCAGACAGATCATTTGAAGTTTACGAAGTTGACATTCCTTGTGGCTCAAGCT 1081
QY 1081 TTTATTTGTTAAGAGACCAACCTTTTACAGAGAGTGTATTAAGAAATTTCCAGAACTTCA 1140
DB 1082 TTTATTTGTTAAGAGACCAACCTTTTACAGAGAGTGTATTAAGAAATTTCCAGAACTTCA 1141
QY 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGAAATTAATGG 1200
DB 1142 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGAAATTAATGG 1201
QY 1201 GAGATTAAGCTTGGTCTTGAAGAGGGACTGTGTCACCTGTTATAGGCAAAAGTTGATGA 1260
DB 1202 GAGATTAAGCTTGGTCTTGAAGAGGGACTGTGTCACCTGTTATAGGCAAAAGTTGATGA 1261
QY 1261 TGGCGTATGTTGGGCTTACAGAGAGATTTCAATATAGGAATCCAGCATTTGGT 1320
DB 1262 TGGCGTATGTTGGGCTTACAGAGAGATTTCAATATAGGAATCCAGCATTTGGT 1321
QY 1321 GAGTTGTTTGGAGAAATCAAAATCGTATGACAAATGATGATCTCCGTGACTATGCAAA 1380
DB 1322 GAGTTGTTTGGAGAAATCAAAATCGTATGACAAATGATGATCTCCGTGACTATGCAAA 1381
QY 1381 TTGTTGGAACCTGTTGGGAGGAGTGTCTTTCTTCTAGTTTCAAGACACCAAAAGATTA 1440
DB 1382 TTGTTGGAACCTGTTGGGAGGAGTGTCTTTCTTCTAGTTTCAAGACACCAAAAGATTA 1441
QY 1441 AAATTTAACTCGGGGACTATGATGATGATTCATATGTTTGAATTAAGTTGAAAGAGTG 1500
DB 1442 AAATTTAACTCGGGGACTATGATGATGATTCATATGTTTGAATTAAGTTGAAAGAGTG 1501
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGTCGACATTAAGGCAAGATTTGAGCCGAG 1560
DB 1502 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGTCGACATTAAGGCAAGATTTGAGCCGAG 1561
QY 1561 CATGTGAAGCTAGTGTCTATGACAGGCACTCAGAAAGTTTCTTCCCTCATATACAGAT 1620
DB 1562 CATGTGAAGCTAGTGTCTATGACAGGCACTCAGAAAGTTTCTTCCCTCATATACAGAT 1621
QY 1621 AGAAATCTGGCTGAACCCAGAGATGTGCAAGAGACATGTTTATGTTAGATTCCTGTTGGT 1680
DB 1622 AGAAATCTGGCTGAACCCAGAGATGTGCAAGAGACATGTTTATGTTAGATTCCTGTTGGT 1681
QY 1681 AACAAATGAGCCGCTGATGTTGAGCTGTTGCTTTTATTCAGAAAGCTGTAAGACCTCT 1740
DB 1682 AACAAATGAGCCGCTGATGTTGAGCTGTTGCTTTTATTCAGAAAGCTGTAAGACCTCT 1741
QY 1741 GAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGGTCTCAGAGAGTACGCTTGAAT 1800
```

```

Db      1742  GAAACCTTTGAACTAATGATTAATGCAATTCAGACTGGGGCTCAGAGAGAGGCTGAT 1801
Qy      1801  GAAACTACTCTGTTGAAAGTCCGTTGCTGATATGTTAAAGAGGCAACTGGAAGATCTTA 1860
Db      1802  GAAACTACTCTGTTGAAAGTCCGTTGCTGATATGTTAAAGAGGCAACTGGAAGATCTTA 1861
Qy      1861  GCTGCTGCTGCTGCAATTTGCACTGATTTCACTGTTGACCCAGAAAGTATTTTCTTAAAGC 1920
Db      1862  GCTGCTGCTGCTGCAATTTGCACTGATTTCACTGTTGACCCAGAAAGTATTTTCTTAAAGC 1921
Qy      1921  AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACCATTA 1980
Db      1922  AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACCATTA 1981
Qy      1981  GGGTCAGTCAAGAGTTCAGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGAG 2040
Db      1982  GGGTCAGTCAAGAGTTCAGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGCAGAG 2041
Qy      2041  AATATGATATCCAAAGTGGCAGAAAGATTAAGTCTGCTTTTGGGCTGATCACCAGATA 2100
Db      2042  AATATGATATCCAAAGTGGCAGAAAGATTAAGTCTGCTTTTGGGCTGATCACCAGATA 2101
Qy      2101  GAAATGTTACCAAGAGTTTGTGATGGCGCAATGCTGAAGATTTGCACTGACAGACAGCT 2160
Db      2102  GAAATGTTACCAAGAGTTTGTGATGGCGCAATGCTGAAGATTTGCACTGACAGACAGCT 2161
Qy      2161  GAAACTGCGCAGCTGGGCTGTTTATGATTAATACATCTGTTGAAACTATCTGTTGACAGT 2220
Db      2162  GAAACTGCGCAGCTGGGCTGTTTATGATTAATACATCTGTTGAAACTATCTGTTGACAGT 2221
Qy      2221  GTGACAGTCTCAGACAGATGGAACCCGCTCTGCTGGAACCACTCTGAGAGAGTCTGCT 2280
Db      2222  GTGACAGTCTCAGACAGATGGAACCCGCTCTGCTGGAACCACTCTGAGAGAGTCTGCT 2281
Qy      2281  TGTCTATCTGATTTGGTTCAATCCAGAAACATGCTACTGATGTCAGAACCTACACACACA 2340
Db      2282  TGTCTATCTGATTTGGTTCAATCCAGAAACATGCTACTGATGTCAGAACCTACACACACA 2341
Qy      2341  AGATACGAAGTTTCTGCTGCAAGTCAAGGCTGGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Db      2342  AGATACGAAGTTTCTGCTGCAAGTCAAGGCTGGAAGAAATCACTGAAGGCTCTGTTCTTGA 2401
Qy      2401  TCATTA 2406
Db      2402  TCATTA 2407

```

```

RESULT 2
LOCUS   AY150490                      2436 bp    mRNA    linear    PLN 23-SEP-2002
DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete
ACCESSION AY150490
VERSION   AY150490.1 GI:23297711
KEYWORDS  FLU CDNA.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
           Burkaya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
           rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2436)
AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
           Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C.,
           Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
           Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
           Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,M.,
           Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
           Ecker,J.R. and Theologis,A.
TITLE     Arabidopsis Open Reading Frame (ORF) clones
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 2436)
AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,

```

```

Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shim,M., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: "RIKEN
Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,
Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shim,M., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
Location/Qualifiers
1..2436
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U13935"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..2436
/gene="At5g42480"
1..2406
/gene="At5g42480"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="PANI2907.1"
/db_xref="GI:23297712"
/translation="MEALSHVIGLSPPOLCRPAPATTKLRSHNTSTTICGASKAD
RLSDPNFTSDSSSFATATTATLVSPSPSIDRBRVPIPIDFYOVGATHTPLT
DGIRAFARVSKPPQFGRSDALLISRLQLOAACTLSPRSRRRENGILDDSEAT
VITDVPMKVPKALCVLQSGEITELVLRGEBALRKLRPSFKQDVLVMAALFLVDS
RAMMALDPDPFTGVEFEVBEALQLQEBGASLAPLRQIDITLEBITPRYVLELGG
LPAGDDYAKRLNGLSGVNRIILMSVGGGASALVGLTRKFNBAFLMPTAAEOVDL
FVATPSNIIPAESFEVYVALVAQAIFGKPKPLQDADKQFQOLQOAKYMAEIPAM
LYDTRNNWEIDRELGICALLIGKDEKRMVIGLSDSOVGNPIVIVEVNSNRD
DNDUDELKLETLALAGYVFRPRTKPKKFLGSPYDDPMLSTLEREVQSGPRL
AAAMARLIGAEHVKASAMQALQKVPSPRTDENSAPROVOSTVSPDVGNGVD
GEGGVPIAIVRSENFETNDYAIRAGVESSEVDETVMSVADMLKASVKLLAAGV
ALGSLFPOKFEFLKSSSFORDKWSMESDVATIGSVRADDSBALPRMDATANRI
VSKWCKISLAFQPDHRIEMLPEVDGRMLKWTDRAAETAGLGVYDITLKLSDVS
VTVASGCTALVATLEESACLSDLVHPNNMTDVRTYTTTRVEVFSKSGWKITBSY
LAS"
2407..2436
/gene="At5g42480"
2433..2434
/gene="At5g42480"
/note="compared to genomic sequence"
/replace="ctg"

```

ORIGIN

```

Query Match          99.7%; Score 2399.6; DB 15; Length 2436;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGAGAGCTGTGATGACGCGGCAATTGCTCTCCCAATTCGAATTATAGCCGATTACCA 60
Db 1 ATGGAAGCTCTGTGATCAGTCGGCATTTGGTCTCTCCCAATTCGAATTATAGCCGATTACCA 60
QY 61 CCGGCGACGACAAAGCTCCGACCTAGCCACAAACCTCTTACAACTATCTGCTCCGCAAG 120
Db 61 CCGGCGACGACAAAGCTCCGACCTAGCCACAAACCTCTTACAACTATCTGCTCCGCAAG 120
QY 121 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTTCACTCTCGAATTCCTCTCTCTCC 180
Db 121 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTTCACTCTCGAATTCCTCTCTCTCC 180
QY 181 TTGCGCACCGGCGACACACCGGCACTCTGCTCTGCGACACATCTATTTGATCGTCC 240
Db 181 TTGCGCACCGGCGACACACCGGCACTCTGCTCTGCGACACATCTATTTGATCGTCC 240
QY 241 GAAAGCGACGTCGCCCATTCGCCATTTGATTTCTACCAAGTATTAGAGCTCAAAACATTTTC 300
Db 241 GAAAGCGACGTCGCCCATTCGCCATTTGATTTCTACCAAGTATTAGAGCTCAAAACATTTTC 300
QY 301 TTAAACCGATGGAATCAGAAAGCATTCGAAGCTAGGCTTTCGAAACCGCGCAATTCGGT 360
Db 301 TTAAACCGATGGAATCAGAAAGCATTCGAAGCTAGGCTTTCGAAACCGCGCAATTCGGT 360
QY 361 TTGAGGAGACGCGCTTTTAACTACCGCGGACAGATTCCTCAAGCTGCTTGCGAACCTG 420
Db 361 TTGAGGAGACGCGCTTTTAACTACCGCGGACAGATTCCTCAAGCTGCTTGCGAACCTG 420
QY 421 TCTAATCTCGCTCTAGAAAGAGATCAATGAAGGCTCTTGTATGATGAAAGAGCTTACA 480
Db 421 TCTAATCTCGCTCTAGAAAGAGATCAATGAAGGCTCTTGTATGATGAAAGAGCTTACA 480
QY 481 GTTCATCATGATGTTCTTGGGATTAAGGTTCTTGGGCTCTCTGTGATTTGCAAGAGGT 540
Db 481 GTTCATCATGATGTTCTTGGGATTAAGGTTCTTGGGCTCTCTGTGATTTGCAAGAGGT 540
QY 541 GGTGAGATCTGAGATGATTTCTTGGGTTGTGAGGCTCTGCTTAAAGAGAGGTTGCCTAAG 600
Db 541 GGTGAGATCTGAGATGATTTCTTGGGTTGTGAGGCTCTGCTTAAAGAGAGGTTGCCTAAG 600
QY 601 TCGTTTAAAGCAATGATGTTTATGATGAGCGCTTGCCTTCTCGATGTCGAGGAGAT 660
Db 601 TCGTTTAAAGCAATGATGTTTATGATGAGCGCTTGCCTTCTCGATGTCGAGGAGAT 660
QY 661 GCTATGGCATTTGATTCACCTGATTTTATTAAGTGTATGAGTTGTTGAGGAAGCTTTG 720
Db 661 GCTATGGCATTTGATTCACCTGATTTTATTAAGTGTATGAGTTGTTGAGGAAGCTTTG 720
QY 721 AAGCTTTTAAAGAGAGAGAGCAAGTACGCTTGCAACCGATTTTACGTCGACAAATTTGAT 780
Db 721 AAGCTTTTAAAGAGAGAGAGCAAGTACGCTTGCAACCGATTTTACGTCGACAAATTTGAT 780
QY 781 GAGACTTTGGAAGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGT 840
Db 781 GAGACTTTGGAAGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGT 840
QY 841 GATGATTTACGCTGGAAGAACTTAATGTTTAAAGCGGTGCGGAATTTTGTGCT 900
Db 841 GATGATTTACGCTGGAAGAACTTAATGTTTAAAGCGGTGCGGAATTTTGTGCT 900
QY 901 GTTGGAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCCGTGGAAGTTTATAT 960
Db 901 GTTGGAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCCGTGGAAGTTTATAT 960
QY 961 GAGGCGTTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
Db 961 GAGGCGTTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
QY 1021 AATATTTCCAGACAGATCATTTGAAGTTTACGAAGTTTCACTTGTCTTGTGGCTCAAGCT 1080
Db 1021 AATATTTCCAGACAGATCATTTGAAGTTTACGAAGTTTCACTTGTCTTGTGGCTCAAGCT 1080

QY 1081 TTTATTTGTAGAAAGCCACCTTTTACAGATGCTGATTAACCAATTCGACACTTCAG 1140
Db 1081 TTTATTTGTAGAAAGCCACCTTTTACAGATGCTGATTAACCAATTCGACACTTCAG 1140
QY 1141 CAGGCTAAGGTAAATGGCTTATGAGATTCCTGGAGTGTGTATGATTAACAGAAATATTG 1200
Db 1141 CAGGCTAAGGTAAATGGCTTATGAGATTCCTGGAGTGTGTATGATTAACAGAAATATTG 1200
QY 1201 GAGATGACTTGGCTTATGAAAGGGGACTCTGTGACCTGCTTATAGGCAAGTTGATGAA 1260
Db 1201 GAGATGACTTGGCTTATGAAAGGGGACTCTGTGACCTGCTTATAGGCAAGTTGATGAA 1260
QY 1261 TGCCTATGCTGGCTTAAACAGTGAAGATTCACAAATATGAAATCAAGTATTGTCG 1320
Db 1261 TGCCTATGCTGGCTTAAACAGTGAAGATTCACAAATATGAAATCAAGTATTGTCG 1320
QY 1321 GAGTTGTTTTTGAAGAAATTCAAATGCTGATCAATGATGATCTCCCTGGAATATGCAAA 1380
Db 1321 GAGTTGTTTTTGAAGAAATTCAAATGCTGATGATCAATGATGATCTCCCTGGAATATGCAAA 1380
QY 1381 TTGTTGAAACCTGGTTGCGAGGGGTGCTTCTCTAGGTTGAGAGACACCAAGATTA 1440
Db 1381 TTGTTGAAACCTGGTTGCGAGGGGTGCTTCTCTAGGTTGAGAGACACCAAGATTA 1440
QY 1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 GAGGTATGTTCAAGGTTCTCTTTAGCTGCTGCAACATAGGCAAGATTTGAGAGCCGAG 1560
Db 1501 GAGGTATGTTCAAGGTTCTCTTTAGCTGCTGCAACATAGGCAAGATTTGAGAGCCGAG 1560
QY 1561 CATGTGAAAGCTAGTCTATGACGAGCACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
Db 1561 CATGTGAAAGCTAGTCTATGACGAGCACTGCAAGAAATTTTCTCCGCTATACAGAT 1620
QY 1621 AGAACTCTGGCTGAACCAAGATGTCAGAGACAGTGTATGATGATGATGATGATGATGATGAT 1680
Db 1621 AGAACTCTGGCTGAACCAAGATGTCAGAGACAGTGTATGATGATGATGATGATGATGATGAT 1680
QY 1681 AACATGTAAGCCGTGATAGTGAAGCTGCTGCTTATATGCAAGACCTGTAAGACCTCT 1740
Db 1681 AACATGTAAGCCGTGATAGTGAAGCTGCTGCTTATATGCAAGACCTGTAAGACCTCT 1740
QY 1741 GAAAACTTTGAAACTAATGATTTATGCAATTCGAGCTGGGGCTCAGAGAGTACGCTTGA 1800
Db 1741 GAAAACTTTGAAACTAATGATTTATGCAATTCGAGCTGGGGCTCAGAGAGTACGCTTGA 1800
QY 1801 GAAACTACTGTTGAAATGTCCTGCTGATATGTTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Db 1801 GAAACTACTGTTGAAATGTCCTGCTGATATGTTTAAAGAGGCAAGTGTGAAGATCTTA 1860
QY 1861 GCTGCTGCTGCTGCAATTTGAGCTGATTTCACTGTTCAGCCAGAAATTTTCTTAAAGC 1920
Db 1861 GCTGCTGCTGCTGCAATTTGAGCTGATTTCACTGTTCAGCCAGAAATTTTCTTAAAGC 1920
QY 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
Db 1921 AGCTCATCTTTTCAAGCAAGATATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
QY 1981 GGGTCAGTGAAGCTGACATTCAGAAAGCACTTCCAGATGAGTCTAGACCTGCAAG 2040
Db 1981 GGGTCAGTGAAGCTGACATTCAGAAAGCACTTCCAGATGAGTCTAGACCTGCAAG 2040
QY 2041 AATATATGATCCAAAGTGGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATCACCCGATA 2100
Db 2041 AATATATGATCCAAAGTGGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATCACCCGATA 2100
QY 2101 GAAATGTTTCCAAAGGTTTTTGAATGGCGAATGCTGAAGATTTGACTGACAGAGAGCT 2160
Db 2101 GAAATGTTTCCAAAGGTTTTTGAATGGCGAATGCTGAAGATTTGACTGACAGAGAGCT 2160
QY 2161 GAAATGCGCAGCTGGGTTGTTTATGATTAATACATGTTGAACATATCTGTTGACAGT 2220

|||||
Db 474 TTCAGGACGACGCTTTAATCAGCCGGAGACAGATTCCTCAAGCTGGTCGAAACCTG 533
421 TCTAACTCCGCTCTGAAGAGAGTCAATGAAGCTCTTGATGATGAAGAAGCTACA 480
534 TCTAACTCCGCTCTGAAGAGAGTCAATGAAGCTCTTGATGATGAAGAAGCTACA 593
481 GTCATCACTGATGTTCTCTGGAGTAAAGTTCCTGGGCTCTCTGTATATGCAAGAAGT 540
594 GTCATCACTGATGTTCTCTGGAGTAAAGTTCCTGGGCTCTCTGTATATGCAAGAAGT 653
541 GGTGAGACTGAGATAGTTCCTGGGCTCTGGAGCTCTGCTTAAAGAGAGTGGCCCTAG 600
654 GGTGAGACTGAGATAGTTCCTGGGCTCTGGAGCTCTGCTTAAAGAGAGTGGCCCTAG 713
601 TCGTTTAAAGCAAAATGCTGTTTAAATGATGATGATGATGATGATGATGATGATGAT 660
714 TCGTTTAAAGCAAAATGCTGTTTAAATGATGATGATGATGATGATGATGATGATGAT 773
661 GCTATGACATTTGATCCACTGATTTTAACTGATGATGATGATGATGATGATGATGAT 720
774 GCTATGACATTTGATCCACTGATTTTAACTGATGATGATGATGATGATGATGATGAT 833
721 AAGCTTTTACAG 780
834 AAGCTTTTACAG 893
781 GAGACTTTGAG 840
894 GAGACTTTGAG 953
841 GATGATTAACGCTCGAAG 900
954 GATGATTAACGCTCGAAG 1013
901 GTTGAAG 960
1014 GTTGAAG 1073
961 GAGCGCTTTTACAG 1020
1074 GAGCGCTTTTACAG 1133
1021 AATATTCAG 1080
1134 AATATTCAG 1193
1081 TTTATTTGTAAG 1140
1194 TTTATTTGTAAG 1253
1141 CAGGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGG 1200
1254 CAGGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGG 1213
1201 GAGATTAAGCTCGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAAT 1260
1314 GAGATTAAGCTCGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAAT 1373
1261 TGCCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGG 1320
1374 TGCCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGG 1433
1321 GAGTTTGTGTTTGAAGATTCAGATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAAT 1380
1434 GAGTTTGTGTTTGAAGATTCAGATGCTTAAGGTAATGCTTAAGGTAATGCTTAAGGTAAT 1493
1381 TTGTTGAAACCTGCTTGAAG 1440
1494 TTGTTGAAACCTGCTTGAAG 1553
1441 AAATTTAAACCTGGGAGCTACTATGATGATGATGATGATGATGATGATGATGATGATGAT 1500

Db 1554 AAATTTAAACCTGGGAGCTACTATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
1501 GAGGTAGTTCAAGGTTCTCTTTAGTGTGTCAGCAATGAGCAAGATGAGAGAGAGAGAG 1560
1614 GAGGTAGTTCAAGGTTCTCTTTAGTGTGTCAGCAATGAGCAAGATGAGAGAGAGAGAG 1673
1561 CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
1674 CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1733
1621 AGAACTGGGCTGAAGCCCAAGAGATGCAAGAGACAGTGTGTTAGTATGATCCTGTTGAT 1680
1734 AGAACTGGGCTGAAGCCCAAGAGATGCAAGAGACAGTGTGTTAGTATGATCCTGTTGAT 1793
1681 AACATGTAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1794 AACATGTAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
1741 GAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAGAGTACCTTGAT 1800
1854 GAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAGAGTACCTTGAT 1913
1801 GAACTTACCTGTTGAAATGTCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1860
1914 GAACTTACCTGTTGAAATGTCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1973
1861 GCTGCTGTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1974 GCTGCTGTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033
1921 AGCTCATCTTTTCAAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
2034 AGCTCATCTTTTCAAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2093
1981 GGTGCTGCTGAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2094 GGTGCTGCTGAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2153
2041 AATATGTATCCAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2154 AATATGTATCCAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2213
2101 GAAATGTATCCAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2214 GAAATGTATCCAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2273
2161 GAAACTGCGCAGCTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
2274 GAAACTGCGCAGCTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2333
2221 GTGACAGTCTCAGCAGATGGAACCGTGTCTGTGTGAAGCAACTGTGAGAGAGTGTGCT 2280
2334 GTGACAGTCTCAGCAGATGGAACCGTGTCTGTGTGAAGCAACTGTGAGAGAGTGTGCT 2293
2281 TGTCTATCTGATTTGGTTTCAATCCAGAAACATGCTACTGATGTCAGAACTTACACACAC 2340
2394 TGTCTATCTGATTTGGTTTCAATCCAGAAACATGCTACTGATGTCAGAACTTACACACAC 2453
2341 AGATACGAAGTTTCTGTGTCAGAGTCAAGGTTGAAGAAATCACTGAAGAGCTGTGTTCTGCA 2400
2454 AGATACGAAGTTTCTGTGTCAGAGTCAAGGTTGAAGAAATCACTGAAGAGCTGTGTTCTGCA 2513
2401 TCATTA 2406
2514 TCATTA 2519

RESULT 4
AY221468 3668 bp DNA linear pln 05-AUG-2003
LOCUS AY221468
DEFINITION Arabidopsis thaliana division protein (AR6) gene, complete cds;
nuclear gene for chloroplast product.
ACCESSION AY221468

```

VERSION      AY221468.1  GI:33436338
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
               Arabidopsis thaliana
               Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
REFERENCE
AUTHORS      1 (bases 1 to 3668)
               Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Erp,H. and
               Osteryoung,K.W.
TITLE         ARC6 Is a U-Domain Plastid Division Protein and an Evolutionary
JOURNAL       Descendant of the Cyanobacterial Cell Division Protein Ftn2
PUBMED        Plant Cell 15 (8), 1918-1933 (2003)
12897262
REFERENCE
AUTHORS      2 (bases 1 to 3668)
               Vitha,S., Koksharova,O., van Erp,H., Froehlich,J.E. and
               Osteryoung,K.W.
TITLE         Direct Submission
JOURNAL       Submitted (17-JAN-2003) Department of Plant Biology, Michigan State
PUBMED        University, 166 Plant Biology Building, East Lansing, MI 48824, USA
12897262
FEATURES
source
               1..3668
               /organism="Arabidopsis thaliana"
               /mol_type="genomic DNA"
               /db_xref="taxon:3702"
               /ecotype="Wassilewskija"
               478..3515
               /gene="ARC6"
               /note="synonym: At5g42480"
               join(478..984,1067..1291,1379..1639,1723..2710,
               3000..3133,3225..3515)
               /gene="ARC6"
               /product="division protein"
               join(478..984,1067..1291,1379..1639,1723..2710,3000..3133,
               3225..3515)
               /note="inner envelope membrane-localized; similar to
               cyanobacterial cell division protein sl10169 of
               Synechocystis; contains U-domain"
               /codon_start=1
               /product="division protein"
               /protein_id="AAQ18645.1"
               /db_xref="GI:33436339"
               /translation="MEALSHVIGLSLPOLCRUPATYKLRSHSTSTTCSKMD
               RILSDPNTSSSSSPATRTTATLVSLDPSIDRBRHPIIDYQVGAATHLT
               DGRFAEAKVSKPQGFSDDLISRQQLQAACETLSRKRREINBLSDEBAT
               VITDVPMDKVPALCVLQEGETEIVLRGELLKRLPKSFQDVLVWALFLVS
               RDMALDPDPTIGYFVEFEBALKLQEGSSSLAPLRAQIDLEBEITPRVYLELG
               LPLGDDYAAKRLNGSGVRNLIWSVGSGSALVGLTRKPMNBAFLRMTAEQVLD
               FVATPSNIPAESFEVYEVALVAQAFLCKPHLLQDADKQPOQLQAQYMAEIRAM
               LYTRNNMEIDPFLERGLCALLIGKVDCEKMLGLDSEDSQYRPAIVEFVLENSRD
               DNDLPGLCLLETWLAGVFPFRFDTKDKKFLGDYDDPMVLSLEREVVQSGPL
               AAATMARIGAEHVKAAMQALOKVPSRTDRNSABPKOVQETFSVDPVGNVRD
               GEGGVFLAEVRSNPENTDVAIRAGVSSSVDETVEVSADMLKEASVKLLAGV
               AILISLFSOKYIKSSSFQRKDMVSMEDVATIGSVADSDALPRMDARTANI
               VSMOKIKSLAPGPDRIEMLPVLDGRMKITWTPDRAEAQGLVYDDTLGLSDS
               VTYSADGTRALVEATLEESACLSDLVHPENNATDVKITTRIEVFMSGSKITBSGV
               LMS"
ORIGIN
Query Match      71.7%; Score 1724; DB 15; Length 3668;
Best local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY      1  ATGAAGCTCTGAGTCAGTCGCGCATTTGCTCTCCCATTCGAATTATGCGGATTAACA 60
DB      478  ATGAAGCTCTGAGTCAGTCGCGCATTTGCTCTCCCATTCGAATTATGCGGATTAACA 537
QY      61  CCGGCGACGACGAAGCTCCGACGTAAGCACAACACTCTTACACTATCTCTCCGCGACG 120
DB      538  CCGGCGACGACGAAGCTCCGACGTAAGCACAACACTCTTACACTATCTCTCCGCGACG 597
QY      121  AATGAGCGACGCGTCTTCTCTCGCACTTCAATTTCACCTCGCATTCCTCTCTCTCC 180

```

```

DB      598  AATGAGCGACGCGTCTTCTCTCGCACTTCAATTTCACCTCGCATTCCTCTCTCTCC 657
QY      181  TTGCGCACCGCCACACACACCGGCACTGTGTCTCTTGCACACCTTATGATCGTCC 240
DB      658  TTGCGCACCGCCACACACACCGGCACTGTGTCTCTTGCACACCTTATGATCGTCC 717
QY      241  GAAAGCGACGTCGCCATCCCATTTGATTTTACAGGATTTAGAGAGCTGAACCATTTG 300
DB      718  GAAAGCGACGTCGCCATCCCATTTGATTTTACAGGATTTAGAGAGCTGAACCATTTG 777
QY      301  TTAACCGATGAATCAAGAGACATTCGAAGCTTGAAGGCTTGAACCGCGCAATTCGT 360
DB      778  TTAACCGATGAATCAAGAGACATTCGAAGGCTTGAAGGCTTGAACCGCGCAATTCGT 837
QY      361  TTCAAGCGACGACGCTTATACAGCGGACAGATTTCTTAACGCTGTTCCGAACCTCTG 420
DB      838  TTCAAGCGACGACGCTTATACAGCGGACAGATTTCTTAACGCTGTTCCGAACCTCTG 897
QY      421  TCTAATCCTCGCTGAAGAGAGTCAATGAAGGCTTCTTGATGATGAAGAGCTTACA 480
DB      898  TCTAATCCTCGCTGAAGAGAGTCAATGAAGGCTTCTTGATGATGAAGAGCTTACA 957
QY      481  GTCATCACTGATGTTCTTGGGAT-----AGGTCCTCGGCG 518
DB      958  GTCATCACTGATGTTCTTGGGATTAAGTAAATTTGATTTGGAATTAAGTTTCTTC 1017
QY      505  -----AGGTCCTCGGCG 518
DB      1018  GTTTTAAATTCATGAATTTGATTAAGAGAACTTTTATCTATGATGAAGGTTCTTCG 1077
QY      519  TCTCTGTGATTCGAAGAGGCTGAGACTGAGATAGTTCTTCGGGTTGAGAGCTCT 578
DB      1078  TCTCTGTGATTCGAAGAGGCTGAGACTGAGATAGTTCTTCGGGTTGAGAGCTCT 1137
QY      579  GCTTAAGAGAGGCTTCTAAGCTGTTTAAAGCAAGATGCTGTTTAAAGCGCTTGC 638
DB      1138  GCTTAAGAGAGGCTTCTAAGCTGTTTAAAGCAAGATGCTGTTTAAAGCGCTTGC 1197
QY      639  GTTTCGTGATGTTCTGAGGAGTCTATGCGATTGATCCACCTGATTTTATCTGTTA 698
DB      1198  GTTTCGTGATGTTCTGAGGAGTCTATGCGATTGATCCACCTGATTTTATCTGTTA 1257
QY      699  TGAAGTTGTTGAGAGGCTTGAAGCTTTAC----- 730
DB      1258  TGAAGTTGTTGAGAGGCTTGAAGCTTTTACAGTAGTTGACTTGTGTAATTG 1317
QY      731  -----A 731
DB      1318  ACGAGCGTTGCTTATTAAGAACTTCTTGATTTGATTACTTTGTTATGAGCTTGTTG 1377
QY      732  GAGAGAGAGCAAGTAGCTTTCGACCGGATTTACGTGACAAATTTGATGAGACTTTG 791
DB      1378  GAGAGAGAGCAAGTAGCTTTCGACCGGATTTACGTGACAAATTTGATGAGACTTTG 1437
QY      792  AAGATCATCTCCGCTTATGCTTGAAGCTAATTCGCTTACCGCTTGATGATTAACG 851
DB      1438  AAGATCATCTCCGCTTATGCTTGAAGCTAATTCGCTTACCGCTTGATGATTAACG 1497
QY      852  TGGGAAAAGCTAAATGTTTAAAGCGTGTGCGGAATATTTTGTCTGTTGAGAGAG 911
DB      1498  TGGGAAAAGCTAAATGTTTAAAGCGTGTGCGGAATATTTTGTCTGTTGAGAGAG 1557
QY      912  TGGAGCATGAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTATGAATGAGCGCTTT 971
DB      1558  TGGAGCATGAGCTCTTGTGGGGGTTTGAACCGGTGAAGATTATGAATGAGCGCTTT 1617
QY      972  ACGAATGACAGCTGCTGAGC----- 991
DB      1618  ACGAATGACAGCTGCTGAGCAGGTATACATTAGATACCTTTTAAATTTCTTTAGCA 1677
QY      992  -----AGGTTGATCTTTTGTGA 1008

```

Db	1678	TGATATACTTTAGAGTTTCTCATTTTAATGTATGTGTGTGTAGTTGATCTTTTGTA	1737
Qy	1009	GCTACCCCCAGCAATATTCCAGCAGAGTCATTTTGAAGTTTACGAAGTTGCACTTGCTCTT	1068
Db	1738	GCTACCCCAACAATATTCAGCAGAGTCATTTGAAGTTTACGAAGTTGCACTTGCTCTT	1797
Qy	1069	GTGGCTCAAGCTTTTATTTGTGTAAAGGCAACCTTTTACAGAGTGTGATTAACCAATTC	1128
Db	1798	GTGGCTCAAGCTTTTATTTGTGTAAAGGCAACCTTTTACAGAGTGTGATTAACCAATTC	1857
Qy	1129	CAGCAACTTCAGCAGGCTTAAGTAAATGCGCTATGAGATTCCTGCGAATGTTGTATGATACA	1188
Db	1858	CAGCAACTTCAGCAGGCTTAAGTAAATGCGCTATGAGATTCCTGCGAATGTTGTATGATACA	1917
Qy	1189	CGAATAATTTGGAGATTAAGTACTGGGTCTAGAAAAGGGACCTGTGTGACCTGCTTAAGGC	1248
Db	1918	CGAATAATTTGGAGATTAAGTACTGGGTCTAGAAAAGGGACCTGTGTGACCTGCTTAAGGC	1977
Qy	1249	AAAGTTGATGAATGCCATATGTGTGTGGGCTTAAACAGTGAAGATTCACATATATAGAAAT	1308
Db	1978	AAAGTTGATGAATGCCATATGTGTGTGGGCTTAAACAGTGAAGATTCACATATATAGAAAT	2037
Qy	1309	CCAGCTATTTGGAGATTTGTTTGGAGAAATCAAAATCGTATGACAAATGATGATCTCCCT	1368
Db	2038	CCAGCTATTTGGAGATTTGTTTGGAGAAATCAAAATCGTATGACAAATGATGATCTCCCT	2097
Qy	1369	GGACTATSCAAATTTGTTGGAAAACCTGTGTGGCAGAGGGTGTCTTCTTCTAGGTTCAAGAC	1428
Db	2098	GGACTATSCAAATTTGTTGGAAAACCTGTGTGGCAGAGGGTGTCTTCTTCTAGGTTCAAGAC	2157
Qy	1429	ACCAAGATTAATAAATTTTAACTCGGGGACATATATGATTCCTATAGGTTTGAATTAC	1488
Db	2158	ACCAAGATTAATAAATTTTAACTCGGGGACATATATGATTCCTATAGGTTTGAATTAC	2217
Qy	1489	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTCCTTTAGCTGCTGCTGCACATATGGCAAG	1548
Db	2218	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTCCTTTAGCTGCTGCTGCACATATGGCAAG	2277
Qy	1549	ATTGAGCCGAGCATGTGAAAGCTAAGTGTCTATGCAAGCATGTCAGAAAAGTTTTCCTTCC	1608
Db	2278	ATTGAGCCGAGCATGTGAAAGCTAAGTGTCTATGCAAGCATGTCAGAAAAGTTTTCCTTCC	2337
Qy	1609	CGCTATACAGATAGAACTCGGCTGTAACCCAGAGATGTGTCAAGAGACAGTGTTAAGTGA	1668
Db	2338	CGCTATACAGATAGAACTCGGCTGTAACCCAGAGATGTGTCAAGAGACAGTGTTAAGTGA	2397
Qy	1669	GATCCGTGTGGTAACAATGTAGGGCCGTATGGTGTGACCTGGAGTCTTATTTGCAAGACT	1728
Db	2398	GATCCGTGTGGTAACAATGTAGGGCCGTATGGTGTGACCTGGAGTCTTATTTGCAAGACT	2457
Qy	1729	GTAAGACCCTCTGAAAACCTTGGAACTATATGATTTATGCAATTCGAGCTGGGGTCTCAGAG	1788
Db	2458	GTAAGACCCTCTGAAAACCTTGGAACTATATGATTTATGCAATTCGAGCTGGGGTCTCAGAG	2517
Qy	1789	AGTAGCGTTGATGAAACTATCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	1848
Db	2518	AGTAGCGTTGATGAAACTATCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	2577
Qy	1849	GTAAGATATCCATAGCTGCTGTGTGTGCAATTTGGAATTCATCTGTGACGACGAAGAT	1908
Db	2578	GTAAGATATCCATAGCTGCTGTGTGTGCAATTTGGAATTCATCTGTGACGACGAAGAT	2637
Qy	1909	TTTCTTAATAAGACGCTCATCTTTTCAAGCAGCATATGTTTCTTCTATGGAATCTGAT	1968
Db	2638	TTTCTTAATAAGACGCTCATCTTTTCAAGCAGCATATGTTTCTTCTATGGAATCTGAT	2697
Qy	1969	GTGCGTACCA-----	1978
Db	2698	GTGCGTACCATATGATGATTTAAATGATGCAATTTTCATATATCTGCAATGCTCAATAATA	2757
Qy	1979	-----	1978
Db	2758	TGCTTGTGTTGTAGCTTAGAACAATAGTTCCCATTTAATACATGTGCCAAAAGTTGTAC	2817

OY		1979	-----	-	1978
Dd		2818	AAGATTAAACAATTTGCTGAATAAATTCACATAATTATGTCTGGAAATTTTTGCATCAA		2877
OY		1979	-----	-	1978
Dd		2878	CTGTAGA CAGAAA TGTAA ATTTCA CTGCA AATTCT GTTTA GAAATTA AGTAGA TTAG		2937
OY		1979	-----	T	1979
Dd		2938	AGATTGCCATTAGTGCGCTTTGCCAACATTTTCTTCCTGATTTTCTTTTCCATTT		2997
OY		1960	AGGCTCAGTCAGACTGACCTACATTCAGAGAACCCTCCAGAA TGGA TCTAGAG ACTGACA		2039
Dd		2998	AGGGTCAGTCAGACTGACCATTTAGAGCACCTTCCA GAAT GGATCTTGAGACTGCAGA		3057
OY		2040	GAAATATAGTATCCAAAGTCGACAGAAATTTAAGTCTTGCTGGCTTTTGGGCTCTATCA CGCAT		2099
Dd		3058	GAAATATAGTATCCAAAGTCGACAGAAATTTAAGTCTTGCTGGCTTTTGGGCTCTATCA CGCAT		3117
OY		2100	AGAAATGTTACCAAG-----		2113
Dd		3118	AGAAATGTTACCAAGAGTGAGAGGAATAAATCTAACATTAATCAATTTGTGTGA AACGTG		3177
OY		2114	-----	-AGTTTGGATGGG-	2128
Dd		3178	TGACATGATTA TAGTCTGCTGCTTGTGGATTTCTGTATTATATAGATTTTGATGGGC		3237
OY		2129	GAAATGCTGAAGATTTTGACATGACACAGACAGCTGMAAC TGGCAGCTTGGTTGTTATG		2188
Dd		3228	GAAATGCTGAAGATTTTGACATGACACAGACAGCTGMAAC TGGCAGCTTGGTTGTTATG		3297
OY		2189	ATTATPACCTGTGAAAACTATCTGTTCGACAGTGTGACAGTCTCAGCGAATGA ACCGTG		2248
Dd		3298	ATTATPACCTGTGAAAACTATCTGTTCGACAGTGTGACAGTCTCAGCGAATGA ACCGTG		3357
OY		2249	CTTGTGTGGAAGCAACTCTGGAGAGATCTGTTGTCATGCTGATTTGGTTCATCCAGAA		2308
Dd		3358	CTTGTGTGGAAGCAACTCTGGAGAGATCTGTTGTCATGCTGATTTGGTTCATCCAGAA		3417
OY		2309	ACAATGCTACATGATGTGCAACTCACAACAACAAGATAGCAAGTTTTCGTCCAGTCCAG		2368
Dd		3418	ACAATGCTACATGATGTGCAACTCACAACAACAAGATAGCAAGTTTTCGTCCAGTCCAG		3477
OY		2369	GGTGGAANAATCAGTGAAAGCTCTGTCTTCCTGCATCATAA		2406
Dd		3478	GGTGAANAATCAGTGAAAGCTCTGTCTTCCTGCATCATAA		3515
RESULT 5 LOCUS DEFINITION ACCESSION VERSION KEYWORDS ORGANISM SOURCE REFERENCE AUTHORS		AY221467 Arbidopsis thaliana truncated division protein (arc6) gene, arc6-1 allele, complete cds; nuclear gene for chloroplast product. AY221467 AY221467.1 GI:33436274 . Arabidopsis thaliana (chale crese) Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosidae; eurosoid II; Brassicales; Brassicaceae; Arabidopsie. 1 (bases 1 to 3664) Vitica.S., Froehlich,J.E., Kokeharova,O., Pyke,K.A., Van Erp,H. and Osteryoung,K.W.	3664 bp DNA linear PLN 05-AUG-2003 arabidopis thaliana truncated division protein (arc6) gene, arc6-1 allele, complete cds; nuclear gene for chloroplast product. GI:33436274		JOURNAL PUMED 12897262 2 (bases 1 to 3664) Vitica,S., Kokeharova,O., van Erp,H., Froehlich,J.E. and Osteryoung,K.W.

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
Source Location/Qualifiers

Source

1. 3664
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Ws-2"
/note="genotype: arc6 mutant"

Gene

1. 3664
/gene="arc6"
/allele="arc6-1"
join(<478..984,1067..1291,1379..>1621)
/gene="arc6-1"
/allele="arc6-1"

mRNA

1. 3664
/product="truncated division protein"
join(478..984,1067..1291,1379..1621)
/gene="arc6"
/allele="arc6-1"

CDS

/note="truncation caused by premature stop codon"
/codon_start=1
/product="truncated division protein"
/protein_id="AAQ18644.1"
/db_xref="GI:33436275"

/translation="MRALSHVIGLSPFOCLRPATYKLRSHNTSTTCSASKRAD
RLSDPNTSDSSSPATKATTTATLVSLPSTIDRPHVPIIDFQVGAQTHFRT
DGIRAEARVSKPQPFSDALISRQLQAACETLSNPSRREBELDEERT
VITVPMKRVGALCVLQEGETEIVLRVEBALKEKLPSPFQDVLVVALFLDVS
RDMAALDPFITGFEVEBALKILQEGESLAPDLRAQIDTLEETITRVYVELLG
LPIGDYAAARLNLGSEVRNIIWSVGGASALVGLTRKKFNMAFL"

ORIGIN

Query Match 71.5%; Score 1720.8; DB 15; Length 3664;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

1 ATGAAGCTCTGATGTCACGTCGGCATTTGTCCTCCCATTCCTCAATATGCCATTACCA 60
478 ATGAAGCTCTGATGTCACGTCGGCATTTGTCCTCCCATTCCTCAATATGCCATTACCA 537
61 CCGGCGACGACAAAGCTCCGACGTAGGCAACAACCTCTCAACTATCTGCTCGGACG 120
538 CCGGCGACGACAAAGCTCCGACGTAGGCAACAACCTCTCAACTATCTGCTCGGACG 597
121 AAATGGGCGACGCTCTCTCTCGACTTCATTTCACTCCGATTCCTCTCTCTCTCC 180
598 AAATGGGCGACGCTCTCTCTCGACTTCATTTCACTCCGATTCCTCTCTCTCTCC 657
181 TTGGCCACCGCCACACACACCGCCACTCTGCTCTGCGACCATTTATGATGCTGCC 240
658 TTGGCCACCGCCACACACACCGCCACTCTGCTCTGCGACCATTTATGATGCTGCC 717
241 GAAGCGACGTCCTCCCATCCCATTTGATTTACACAGATTTAGGAGCTCAACAACATTTCC 300
718 GAAGCGACGTCCTCCCATCCCATTTGATTTACACAGATTTAGGAGCTCAACAACATTTCC 777
301 TTAAACGATGAATTCAGAAAGATTCGAAGCTTAGGTTTCGAAACCGCGCAATTCGCT 360
778 TTAAACGATGAATTCAGAAAGATTCGAAGCTTAGGTTTCGAAACCGCGCAATTCGCT 837
361 TTCAAGGACGACGCTTTAATCAGCCGAGACACATTTTCAAGCTGTTGCGAAACTCTG 420
838 TTCAAGGACGACGCTTTAATCAGCCGAGACACATTTTCAAGCTGTTGCGAAACTCTG 897
421 TCTAATCTCTGCTAGAGAGATACATGAAGGCTTCTGATGATGAAGAGCTACA 480
898 TCTAATCTCTGCTAGAGAGATACATGAAGGCTTCTGATGATGAAGAGCTACA 957
481 GTCATCACTGATGTTCTTGGGAT-----504
958 GTCATCACTGATGTTCTTGGGATTAAGTATTCGATTTCCGAAATATAAGTTTCTTC 1017

505 -----AGGTCCTGGGCG 518
1018 GTTTTAATTTGATGAATGGAATAAGAGAACTTTATCTAGTGAAGGTTCTTGCGGC 1077
519 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCCGGTGTGAGGCTCT 578
1078 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCCGGTGTGAGGCTCT 1137
579 GCTTAAGAGAGGTGTGAGGTGATGCTTTAAGCAAGATGTGTTTACTTATGCGGCTTGC 638
1138 GCTTAAGAGAGGTGTGAGGTGATGCTTTAAGCAAGATGTGTTTACTTATGCGGCTTGC 1197
639 GTTTCGATGTCGAGGAGTGTGATGCTTGAAGTTCACCTGATTTTATCTGCTGTA 698
1198 GTTTCGATGTCGAGGAGTGTGATGCTTGAAGTTCACCTGATTTTATCTGCTGTA 1257
699 TGAAGTTGTTGAGAAAGCTTTGAAGCTTTTAC-----730
1258 TGAAGTTGTTGAGAAAGCTTTGAAGCTTTTACAGGTAGTTGACTTGTGTAATTTG 1317
731 -----A 731
1318 ACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATTAATTTGATGCTGTGTA 1377
732 GAGAGAAAGAGCAAGTAGCTTGACCGGATTTACGTCACAAATTGATGAGACTTTGGA 791
1378 GAGAGAAAGAGCAAGTAGCTTGACCGGATTTACGTCACAAATTGATGAGACTTTGGA 1437
792 AGAGATCACTCCGCGTTATGCTTTGAGACTACTTGGCTTACCGCTTGTGATGATTAACG 851
1438 AGAGATCACTCCGCGTTATGCTTTGAGACTACTTGGCTTACCGCTTGTGATGATTAACG 1497
852 TCGAAAAGACTTAATGTTTAAAGCGGTGTGCGGAATTTTGTGCTGTGAGAGAGG 911
1498 TCGAAAAGACTTAATGTTTAAAGCGGTGTGCGGAATTTTGTGCTGTGAGAGAGG 1557
912 TGAGCATCAGCTCTGTTGGGGTTTGAACCGGTGAAGTTTATGAATGAGGCGTTTTC 971
1558 TGAGCATCAGCTCTGTTGGGGTTTGAACCGGTGAAGTTTATGAATGAGGCGTTTTC 1617
972 ACGAATGACAGCTGTGAGC-----991
1618 ACGAATGACAGCTGTGAGCAGGATTAACAGTTTATTTTATTTCTTTAGCA 1677
992 -----AGTTGATCTTTTGTGTA 1008
1678 TGAATTAATTTAGTTTCTCATTTTAATGATGTTGTGATGATGATGATGATGATGATGAT 1737
1009 GTCACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGAGTTGCACTTGCTCT 1068
1738 GTCACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGAGTTGCACTTGCTCT 1797
1069 GTGGCTCAAGCTTTTATTTGTAAGAACCAACCTTTTACAGAGTCTGATTAAGCAATTC 1128
1798 GTGGCTCAAGCTTTTATTTGTAAGAACCAACCTTTTACAGAGTCTGATTAAGCAATTC 1857
1129 CAGCAACTTCAGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1188
1858 CAGCAACTTCAGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1917
1189 CCGAATTAATTTGGAGATGAGCTTCCGCTTGAAGAGGAGCTGTCGACCTGTTATAGGC 1248
1918 CCGAATTAATTTGGAGATGAGCTTCCGCTTGAAGAGGAGCTGTCGACCTGTTATAGGC 1977
1249 AAAGTTGATGATGCGGTATGTTGTTGAGGCTTGAAGAGGATTCACAAATATAGGAAT 1308
1978 AAAGTTGATGATGCGGTATGTTGTTGAGGCTTGAAGAGGATTCACAAATATAGGAAT 2037
1309 CCAAGTATTTGAGAGTTTGTGTTGAGGATTCACAAATGATGATGATGATGATGATGATGATGAT 1368
2038 CCAAGTATTTGAGAGTTTGTGTTGAGGATTCACAAATGATGATGATGATGATGATGATGATGAT 2097
1369 GCACTATGCAAAATGTTGGAAGCTGTGTTGAGGAGGTTGTCTTCTAGGTTCAAGAGAC 1428

```

Db      2098  GGAAGTATGCAAAATGTTGAAAACCTGGTGGACAGGGGTTGCTTCTAGGTTGAGAGAC 2157
Qy      1429  ACCAAGATATAAAATTTAAATCTCGGGACTACTATGATATGCTATGTTTGAATTAC 1488
Db      2158  ACCAAGATATAAAATTTAAATCTCGGGACTACTATGATATGCTATGTTTGAATTAC 2217
Qy      1489  TTGAAAGAGTGGAGGTAGTTCAGAGGTTCTCTTTAGCTCTGCTGCAATATGCAAG 1548
Db      2218  TTGAAAGAGTGGAGGTAGTTCAGAGGTTCTCTTTAGCTCTGCTGCAATATGCAAG 2277
Qy      1549  ATTGAGCCGAGCATGTGAAGCTAGTCTATGAGGCACTGCAAGAAATTTTCTTCC 1608
Db      2278  ATTGAGCCGAGCATGTGAAGCTAGTCTATGAGGCACTGCAAGAAATTTTCTTCC 2337
Qy      1609  CGCTATACATATGAAAACCTGGCTGAAACCAAGATGTGCAAGACAGTGTATGTA 1668
Db      2338  CGCTATACATATGAAAACCTGGCTGAAACCAAGATGTGCAAGACAGTGTATGTA 2397
Qy      1669  GATCTGTTGGTAAACAATGATAGCCGTGATGATGAGCTGATGCTTATTTGCAAGCT 1728
Db      2398  GATCTGTTGGTAAACAATGATAGCCGTGATGATGAGCTGATGCTTATTTGCAAGCT 2457
Qy      1729  GTAAGACCTCTGAAAACTTTGAAACTAATGATTATGCAATTCGAGCTGGGCTCTCAG 1788
Db      2458  GTAAGACCTCTGAAAACTTTGAAACTAATGATTATGCAATTCGAGCTGGGCTCTCAG 2517
Qy      1789  AGTAGCCTTATGATAAATCTACTGTTGAAATGCTGCTGCTATATGTTAAAGAGGCA 1848
Db      2518  AGTAGCCTTATGATAAATCTACTGTTGAAATGCTGCTGCTATATGTTAAAGAGGCA 2577
Qy      1849  GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
Db      2578  GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2637
Qy      1909  TTTCTTAAAGACAGCTCATCTTTTCAACGAGAGATATGTTTCTTATGAAATCTGAT 1968
Db      2638  TTTCTTAAAGACAGCTCATCTTTTCAACGAGAGATATGTTTCTTATGAAATCTGAT 2697
Qy      1969  GTGCTTACCA----- 1978
Db      2698  GTGCTTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGATGCTCAATAA 2757
Qy      1979  ----- 1978
Db      2758  TGCTTGTGTTGTAGCTAAGAACATAGTCCCACTTAATACATGCTCCCAAAAGTTGAC 2817
Qy      1979  ----- 1978
Db      2818  AAGATTAACAAGTTGCTGAGTAATTTCACTAATTATGCTGTAATTTTGTGATCAAA 2877
Qy      1979  ----- 1978
Db      2878  CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATGAGATTAAG 2937
Qy      1979  -----T 1979
Db      2938  AGATTGCTTAAGTGTGCTTGTCTCAACTTTTCTTCTGATTTTCTTCTTCAATTT 2997
Qy      1980  AGGGTCAGTCAGAGCTGAGATTTGAGAGCACTTCCAGAAATGAGTCTAGAGCTGAGA 2039
Db      2998  AGGGTCAGTCAGAGCTGAGATTTGAGAGCACTTCCAGAAATGAGTCTAGAGCTGAGA 3057
Qy      2040  GAATATATGATATCCAAAGTGGCAAGAAATTAATGCTCTGCTTGGGCTGATCAACCC 2099
Db      3058  GAATATATGATATCCAAAGTGGCAAGAAATTAATGCTCTGCTTGGGCTGATCAACCC 3117
Qy      2100  AGAATATGATCCAG----- 2113
Db      3118  AGAATATGATCCAGAGGTGAGGGAATAATCTACAAATTCATCAATGTTGTGAAACTGT 3177
Qy      2114  -----AGGTTTGGATGGGC 2128

```

```

Db      3178  TGACATGATATATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3237
Qy      2129  GAATGCTGAAGATTTGAGTCTGACAGAGCACTGAAATCTGCGACGTTGGGTTTATG 2188
Db      3238  GAATGCTGAAGATTTGAGTCTGACAGAGCACTGAAATCTGCGACGTTGGGTTTATG 3297
Qy      2189  ATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCC 2248
Db      3298  ATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCC 3357
Qy      2249  CTCTGTTGAGAACCACTCTGAGAGGAGTCTGTTGCTATCTGATTTGGTTATCCAGAA 2308
Db      3358  CTCTGTTGAGAACCACTCTGAGAGGAGTCTGTTGCTATCTGATTTGGTTATCCAGAA 3417
Qy      2309  ACAATGCTACTGATGTGACAGACCTTACACAAAGATGATGATGTTTCTGGTCCAACT 2368
Db      3418  ACAATGCTACTGATGTGACAGACCTTACACAAAGATGATGATGTTTCTGGTCCAACT 3477
Qy      2369  GGTGAAATCACTGAAAGGCTCTGTTCTTGATCATTA 2406
Db      3478  GGTGAAATCACTGAAAGGCTCTGTTCTTGATCATTA 3515

```

```

RESULT 6
AB016888
LOCUS
DEFINITION
Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MDH9.
ACCESSION
AB016888
VERSION
AB016888.1
KEYWORDS
GT:3449329
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

```

```

REFERENCE
AUTHORS
Asamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kocant, H., Miyajima, N.
and Tabata, S.
TITLE
Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
Sequence features of the regions of 1,081,958 bp covered by
seven physical clones assigned pl and TAC clones
JOURNAL
DNA Res. 5 (6), 379-391 (1998)
PUBMED
10048488
REFERENCE
2 (bases 1 to 85791)
AUTHORS
Nakamura, Y.
TITLE
Direct Substitution
JOURNAL
Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)

```

```

COMMENT
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd.graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Graal
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://greenlii.cool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K514 and the 3' clone is K16E1.
Location/Qualifiers
1..85791
/organism="Arabidopsis thaliana"

```

```

FEATURES
source

```

exon
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MDH9"
/clone_1b="Miteui P1"
/ecotype="Columbia"
complement (59. .239)
/note="CDS is reported in Acc# AB023032
gene_id:KSJ14.12
unknown protein"
/number=1

CDS
/evidence=not experimental
complement (join(2056. .2079,2278. .2393,2991. .3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pit|T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. .6138,6462. .6595)
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1

CDS
/evidence=not experimental
/protein_id="BABI0474.1"
/db_xref="GI:9759469"
/translation="MWDPEAEEDRKKSNLSLAKNSVLPERGRFYEEYAKRE
RLRRKGEAEVGGVKGTPYNGVNEHPMTRKRGYAKESIKRTVSVESSTIR
YSLRSMKNERKPPPLPLNVAVSMKSVTTERKRLSLTCTACTSSGLAFRLTGS
EPGVHAKRTLPFSALV"
complement (join(6868. .7911,7989. .8414,8555. .8671,
8754. .9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="arm repeat containing protein"
/protein_id="BABI0475.1"
/db_xref="GI:9759470"
/translation="MOTDEBATGDABRDEEVYAEERIHBEVDAVEIHEEYKEDD
DCEGLVSDIVSEFLDQINGYRTOQECFNLVRKILIPILDEIRGESPSCKF
LNRKRVFLAKKLLETCSNGSKIYALDGETMTRPSIYEKRLVLAAPDEMI
SGADKEDISLCKQKAKKRTDQDIELAVMMVPSKTPRNASAIIERLAKKLE
LOTIDDKETETIAIOSLIDPKGLNIEIKOHIIELEKPKKQGLBATDITDYOVIK
AIRTSTLIPHEPICITLLEIMDPIYIANGOTYKESIOKMPDAGHKCPTROEL
DHLSAPNPLKULIMQCEKKNFKIPEKVSPOSQOEKDYSLVEALSSSOLEEQ
RRSVQKQRLARENPERVLIANAAGALPLIVQLSYDSGIQENAVTTLNLSDIEN
KCLISNEGALPNIEILENGRREARNSAALFSLMLDENKVTIGLSNGIPLVLDL
OHTLGRKQDALTALFNLSLNSANKRAIDAGIVQPLNLDKNGIDELAILLL
LASHPEGROAIGOLSFLETVEFROSTPRNKCATSVLLELGSNNSSFTLALQFV
YEVLETTTSGTRAKRANALIQILKSEQT"

CDS
10888. .12579
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BABI0476.1"
/db_xref="GI:9759471"

CDS
/translation="MMISEKPLGEBESIRDLBYLTVSRRLVYSVQKLKKIKHTEV
VEDBEIARAVNCLISVGCRAVDGDEPDSSNKRWSASBEKGLMTICGTEETL
DCPSYGRERFMKNNRKYLDAGODRKHYVLPDILEMCLMRPLTSLNHLVYK
KMSMANTQRFLOMRREGSFQTPMLPLPALKQCCSGDTHGVDSDKMRLETDL
KGRPMYSVISIHEEYIVGGRSMDRNSFKSHRGILVSPSIKAMRKASRHAASLPT
VGAETVSESTQTKQNRQDRPHLSRVGSESDYEDPHRLSVRRQNRASAOQNGK
SHRLIROKLDRLNRSKRPVLIAGTGTFDEPLDSGEIYDSATWMSMOQLPMKF
GVVSCGICNGIFAYASENDKLSGYDIERGFMTIGTSPIPRVHIEFYPXLVCSNHL
FMLSVSCDGDQOIGRNRKAVKRMELDLYLTWESVHPADPMDMNTATVSDONI
LMGIEMKITFGQVLSFTYVDUTLEBSMHSVSNQSQCLNLSCTKTITALLHL"
complement (15087. .16778)
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1

CDS
/evidence=not experimental
/protein_id="BABI0477.1"
/db_xref="GI:9759472"
/translation="MMISEKPLGEBESIRDLBYLTVSRRLVYSVQKLKKIKHTEV
EDBEVARGAVNCLISVGCRAVDGDEPDSSNKRWSASBEKGLMTICGTEETL
DCPSYGRERFMKNNRKYLDAGODRKHYVLPDILEMCLMRPLTSLNHLVYK
KMSMANTQRFLOMRREGSFQTPMLPLPALKQCCSGDTHGVDSDKMRLETDL
KGRPMYSVISIHEEYIVGGRSMDRNSFKSHRGILVSPSIKAMRKASRHAASLPT
VGAETVSESTQTKQNRQDRPHLSRVGSESDYEDPHRLSVRRQNRASAOQNGK
SHRLIROKLDRLNRSKRPVLIAGTGTFDEPLDSGEIYDSATWMSMOQLPMKF
GVVSCGICNGIFAYASENDKLSGYDIERGFMTIGTSPIPRVHIEFYPXLVCSNHL
FMLSVSCDGDQOIGRNRKAVKRMELDLYLTWESVHPADPMDMNTATVSDONI
LMGIEMKITFGQVLSFTYVDUTLEBSMHSVSNQSQCLNLSCTKTITALLHL"
complement (join(18111. .18193,18281. .18535,18773. .19001,
19088. .19195,19307. .19447,19542. .19952,20129. .20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
sp|P42251"

CDS
/evidence=not experimental
/protein_id="BABI0478.1"
/db_xref="GI:9759473"
/translation="MMQISFWCLSLTLTVIAAEPTIRIARFSCANOGAPOPIWDA
INKEFDQLFIMLDGNTYGDIRPLTVIGKERTGPMKNSRFPVSSSEKRLRYAKK
ANPGYRIQRKAVIGTMDHDYGLNDAGKEFPRKYINQKLMADPLNEPLDSRKAQ
GVYASYTGPENRKYIYADTRHDPKSDSILGDTQMDMLENLSGPRSEITII
GSSVOYISNLSATQTFYMSWGRPRPKRKLPLADTKRGVPIPSGVHREGIT
KIDCSGYPLIDTSSGLVOSVKKVPRPLRSIVRLFPYTPPTMRYINDCKFSCT
YQOONFALSIDNNANPVTKLEIRVNGDTVIGTVSSELDPOGSSNLSLDTFTKK
SQRYCTLEIEPLPRTYRLAVLIFPIYAVLMAVILGILGAVALITACVYKCRVD"
complement (21540. .22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BABI0479.1"
/db_xref="GI:9759474"
/translation="MTLAKQKSLRLYKVSRSKSESRNLEDESRSSNGSSGS
LNVNEKTFYDNDANSQKISGEBELSCVSLGALSSEVEYKTSVYDDQAPID
FEERLKLMEBESDSEIRKELEAGVMYEBEFTTAAISLRTTSLRDESCTVDAC
KVMIRFQDDNDVLSFDEFVLMNR"
join(24090. .24289,24410. .24614,24701. .24945,25115. .25313,
25632. .25679,25821. .25928,26039. .26227,26314. .26613,
26730. .26831,27169. .27303,27487. .27711,27919. .28122,
28213. .28380,28646. .28795,28889. .29026,29204. .29431,
29519. .29671,29784. .29910,30115. .30241,30554. .30674,
30763. .30843,30932. .31033,31112. .31262,31338. .31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="p1c11ybsn"
/protein_id="BABI0480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFTVGKFSPIILAPFNSGDSRRARYLSDSNKRYFNS
SPILTPHRYVVEAPSLIPYNGILMAOPNSHGKGLKXNYSGKATGISLSQGNFCLT
CRNQGIRALDPSATVDRTAFSLSSSLTSSLRKSQIYNATLGPBEPRAAGAMD
GIVAEKODDLPLPEIDSALEFLGCELPKSHRKLHKGKNGRQIYLPKPPNF
BAHMEAEVDEIDEEDEQIAHMIIEHAFGLSKREGLDAGARSNAVYDTPHHTVPI
HSPLTHKSEDDLPFSLDALNLIIEHPKLSRVEREKRAIISBLQMMNTIEYRPHI
OLLOHLSHKKLGRFPPIGLEBOIKKMDYKIRFHRMYFANATLYIYGDINDPR
IYHNIIEHAKRGLDNESTPSSPSPCAFAQMANPLVPLKAGIGTSPSTKNTAOS
KMTREHARIPPEVHNSIPGTSVULKFPQITKHELQNFALNMEKTIYVQCTG
DLNVMKRIETSLALHRIINTRYKSNNPETSVELHSDSGRCCTVTTLTVABEPN
MONAVVAVAOEVARLKEFGVTRBELTRYMDALKDSBHLAAMDVNSVNDPIMES
DASHVTOQOHEFTLVAVAGVTLEEVTVGAKYLEFISDGRPTALPLAIVCV
PTKTVHVGSESPENISPDILISVSGSLAPLEAPBELEPVELISQSLKELTOL
NPGFVPIPSGILKADKCGITLOLNSCIYANKSTESAPGMRILVGGGRABR
TSDSKAAVVGATTSLEGRAVDPSRQVLEPVMNHLINCSLESTEFIAMERPTLR
DNMQQAFLHLNHLERSWLEDAFDRARQLYSTRSPKSLERKTAHKLAMLNG
DERFVEPTPKSLQSLNLSKYDAVMHSPVDNNNEVSIIVGFSSEIERCLTDLVGLYK
AFSDSKAPGSEBILFROPYAGIOFOVFLKQDDEBAICAIAPAPRMQFTVDDGDL
FQSVSKLPVANDILKSEBQLEGGRELOKRLRAPLPGVTMGLLAIINSLFTT

CDS

VDLSGLTYDVSEFLNLPRLNLGMVTVSVTSIPGVYKAVADACKSVLRLSHNOIAP
RELRAKRTILMKHEALIKSNAYMLNLQASVPRKLSIKELVSLYEASIED
ITLVYQLHVEDSLVSCIGIAGQAGEEITVLSEBEPEDSVGVPGVSGSMTR
PTT"
complement (join(33251..33326,33512..33663,33744..33797,
33888..33968,34143..34285,34359..34768,34873..35052,
35052..35052))

Query Match 71.4%; Score 1717.6; DB 15; Length 85791;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

QY 1 ATGAAAGCTCTGATCAGTCCGCACTTCTCCCATTCATTTATCCGATTACCA 60
Db 64077 ATGAAAGCTCTGATCAGTCCGCACTTCTCCCATTCATTTATCCGATTACCA 64136
QY 61 CCGGCGACGCAAGAGCTCCGACGTAGACCAACACCTCTCACTATCTCCGCGACG 120
Db 64137 CCGGCGACGCAAGAGCTCCGACGTAGACCAACACCTCTCACTATCTCCGCGACG 64196
QY 121 AATGAGGCGACGCTCTCTCCGACTTCTCAATTTCTCCGATTCTCTCTCTCC 180
Db 64197 AATGAGGCGACGCTCTCTCTCCGACTTCTCAATTTCTCCGATTCTCTCTCTCC 64256
QY 181 TTGCGCACCGGCAACACGCGCACTCTCTCTCTCTCTCCACCATCTATTGATCGTCC 240
Db 64257 TTGCGCACCGGCAACACGCGCACTCTCTCTCTCTCCACCATCTATTGATCGTCC 64316
QY 241 GAAAGCGACGCTCCCATCCCATTTGATTCTACAGGATTTAGAGAGTCAACACATTTCC 300
Db 64317 GAAAGCGACGCTCCCATCCCATTTGATTCTACAGGATTTAGAGAGTCAACACATTTCC 64376
QY 301 TTAAACGATGATCAGAAAGAGATTCGAGCTAGGCTTTCGAAACCGCGCAATTCGAT 360
Db 64377 TTAAACGATGATCAGAAAGAGATTCGAGCTAGGCTTTCGAAACCGCGCAATTCGAT 64436
QY 361 TTGAGCGACGACGCTTTAATCAGCCGCGAGACAGATTTCTCAAGCTCTCCGAAACTCTG 420
Db 64437 TTGAGCGACGACGCTTTAATCAGCCGCGAGACAGATTTCTCAAGCTCTCCGAAACTCTG 64496
QY 421 TCTAATCTCTCGCTCTAGAAAGAGATTCAGAAAGCTCTCTCTATGATGTAAGAAAGCTACA 480
Db 64497 TCTAATCTCTCGCTCTAGAAAGAGATTCAGAAAGCTCTCTCTATGATGTAAGAAAGCTACA 64556
QY 481 GTCATCACTGATGTTCTCTGGAGT----- 504
Db 64557 GTCATCACTGATGTTCTCTGGAGT----- 64616
QY 505 -----AAGTTCTCTGGAGC 518
Db 64617 GTTTAATTTCAATGATTTGATAAAGAAAGAACTTTATCTAGTGAAGGTTCTCTGGAGC 64676
QY 519 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATGTTCTCTGGGTGTGAGGCTCT 578
Db 64677 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATGTTCTCTGGGTGTGAGGCTCT 64736
QY 579 GCTTAAGAGAGGTGCTAAGTCTGTTAAGCAAGATGTGTTTATGTTATGCGGCTTGC 638
Db 64737 GCTTAAGAGAGGTGCTAAGTCTGTTAAGCAAGATGTGTTTATGTTATGCGGCTTGC 64796
QY 639 GTTTCCTGATGTCTCGAGGAGTGTGATGATTCACCTGATTTTATATCTGTTA 698
Db 64797 GTTTCCTGATGTCTCGAGGAGTGTGATGATTCACCTGATTTTATATCTGTTA 64856
QY 699 TGAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 730
Db 64857 TGAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 64916
QY 731 -----A 731
Db 64917 ACGAGCCTTGGCTTTAAGAACTTTCTGATTTGATACCTTGTATGAGCTTGTGTA 64976
QY 732 GGAGAGAGAGCAAGTAGCTTGCACCGGATTTACGTGCAAAATGATGAGACTTTTGA 791
Db 732 GGAGAGAGAGCAAGTAGCTTGCACCGGATTTACGTGCAAAATGATGAGACTTTTGA 791

Db 64977 GGAGAGAGAGCAAGTAGCTTGCACCGGATTTACGTGCAAAATGATGAGACTTTTGA 65036
QY 792 AGAGATCACTCCGCGTTATGTCTTGAGAGTACTTGGCTTACCGCTTGTGATGATTACGC 851
Db 65037 AGAGATCACTCCGCGTTATGTCTTGAGAGTACTTGGCTTACCGCTTGTGATGATTACGC 65096
QY 852 TGGCAAAAGACTAATAGTGTAAAGCGGTGTGCGGAATATTTTGTGCTCTGTGAGAGAG 911
Db 65097 TGGCAAAAGACTAATAGTGTAAAGCGGTGTGCGGAATATTTTGTGCTCTGTGAGAGAG 65156
QY 912 TGGAGATCAGCTCTGTTGGGGGTTTGAACCGGTGAGAGATTATGATGAGGCGTTTGT 971
Db 65157 TGGAGATCAGCTCTGTTGGGGGTTTGAACCGGTGAGAGATTATGATGAGGCGTTTGT 65216
QY 972 ACGAATGACAGCTGTGAGC----- 991
Db 65217 ACGAATGACAGCTGTGAGCAGGATATACAGTTAGATCTTTTAAATTTCTTACGCA 65276
QY 992 -----AGTTGATCTTTTGTGA 1008
Db 65277 TGATATACCTTTAGTGTCTCATTTTAAATGATGATGATGATGATGATGATGATGATGAT 65336
QY 1009 GCTACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAGTTGCACTTGTCTT 1068
Db 65337 GCTACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAGTTGCACTTGTCTT 65396
QY 1069 GTGGCTCAAGCTTTTATGTTAAGAGACCACTTTTACAGAGTCTGTATGAACAATTC 1128
Db 65397 GTGGCTCAAGCTTTTATGTTAAGAGACCACTTTTACAGAGTCTGTATGAACAATTC 65456
QY 1129 CAGCAACTTACAGAGCTTAAAGTATGCTATGAGATTCCTGCAATGTTATGATATCA 1188
Db 65457 CAGCAACTTACAGAGCTTAAAGTATGCTATGAGATTCCTGCAATGTTATGATATCA 65516
QY 1189 CGGAATTAATTTGGAGATATGCTCGGCTTGAAGAGGGGACTCTGTCACCTGTTATAGGC 1248
Db 65517 CGGAATTAATTTGGAGATATGCTCGGCTTGAAGAGGGGACTCTGTCACCTGTTATAGGC 65576
QY 1249 AAGTTGATGATATGCGGATATGTTGAGCTTGAAGAGAGATTCACAAATATAGGAAT 1308
Db 65577 AAGTTGATGATATGCGGATATGTTGAGCTTGAAGAGAGATTCACAAATATAGGAAT 65636
QY 1309 CAGCTATTGTGAGATTTGTTTGAAGATTCAAATCTGATGACATGATGATCTCTCT 1368
Db 65637 CAGCTATTGTGAGATTTGTTTGAAGATTCAAATCTGATGACATGATGATCTCTCT 65696
QY 1369 GGAATATGCAAAATTTGTAAGAACTGTTGGAGGGGTTGCTCTCTAGGTTCAAGAG 1428
Db 65697 GGAATATGCAAAATTTGTAAGAACTGTTGGAGGGGTTGCTCTCTAGGTTCAAGAG 65756
QY 1429 ACCAAAGATPAAATTTTAACTCGGGGACTACTATGATGATCTATGTTTGAAGTTAC 1488
Db 65757 ACCAAAGATPAAATTTTAACTCGGGGACTACTATGATGATCTATGTTTGAAGTTAC 65816
QY 1489 TTGAAAAGAGTGAAGTGAAGTTCAAGGTTCTCTTAACTGCTGCTGCACTATGCAAGG 1548
Db 65817 TTGAAAAGAGTGAAGTGAAGTTCAAGGTTCTCTTAACTGCTGCTGCACTATGCAAGG 65876
QY 1549 ATTGAGAGCGAGATGTAAGAACTGATGCTATGCAAGCACTGAGAAAGTTTCTCTCC 1608
Db 65877 ATTGAGAGCGAGATGTAAGAACTGATGCTATGCAAGCACTGAGAAAGTTTCTCTCC 65936
QY 1609 CGCTATACAGATGTAAGAACTGCTGTAACCAAGAGATGTGCAAGAGACAGTGTATGATGA 1668
Db 65937 CGCTATACAGATGTAAGAACTGCTGTAACCAAGAGATGTGCAAGAGACAGTGTATGATGA 65996
QY 1669 GATCTGTGTTGTAACAATGATGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGAGCT 1728
Db 65997 GATCTGTGTTGTAACAATGATGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGAGCT 66056
QY 1729 GTAAGACCTCTGTAACAATTTGAACATAATGATTTAAGCAATGAGAGCTGGGGCTGAGAG 1788
Db 66057 GTAAGACCTCTGTAACAATTTGAACATAATGATTTAAGCAATGAGAGCTGGGGCTGAGAG 66116

QY	1789	GTGAGCGTGGATNGAAACACTGCTGGAATGTCGCTGCTGATATGATGAAAGGAGCGAAGT	1848
Db	66117	AGTACCGTGTATGAAACTACTGTTGAAATGTCCTGTCTGATATGTTAAAGAGGCGAAGT	6617
QY	1849	GTGAAGATCCCTAGCTGCTGTGTGGCAATTGGACTGATATTTCACTGTTCAAGCCAGAAATAT	1908
Db	66177	GTGAAGATCCCTAGCTGCTGTGTGGCAATTGGACTGATATTTCACTGTTCAAGCCAGAAATAT	6623
QY	1909	TTTCTTTAAAGACAGCTCATCTTTTCCAAAGCAGAGATATGCTTTCTTATGAAATCTGAT	1968
Db	66237	TTTCTTTAAAGACAGCTCATCTTTTCCAAAGCAGAGATATGCTTTCTTATGAAATCTGAT	6629
QY	1969	GTGCGTATCA-----	1978
Db	66297	GTGCGTATCAATGATGATTTTCAATATATCTGATTTCTCAAAATAT	6635
QY	1979	-----	1978
Db	66357	TGCTTGTTTTGTGAGCTAAGAACATATGTTCCACTTAATATCATGTGCCAAAGTTGTACC	6641
QY	1979	-----	1978
Db	66417	AAGATTACAAAGTTGCTGAGTAAATTTCACTAATTAATGCTGTGAATTTTTTGATCAAA	6647
QY	1979	-----	1978
Db	66477	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAGAAATACGTAGATTAG	6653
QY	1979	-----	1979
Db	66537	AGATTGCTTAGTGTGCTGTGTGTCACACTTTTCTTCTGATTTTCTTTTCATTT	6659
QY	1980	AGGCTCACTCAGAGCTGACATTTCAAGAACACTTCCCAAGATGATCTTAGACTGACA	2039
Db	66597	AGGCTCACTCAGAGCTGACATTTCAAGAACACTTCCCAAGATGATCTTAGACTGACA	6665
QY	2040	GAATATAGTATCCCAAGTGGCAGAAAGATTAATAGTCTGTGGCTTTTGGGCTGATCAAGCGAT	2099
Db	66657	GAATATAGTATCCCAAGTGGCAGAAAGATTAATAGTCTGTGGCTTTTGGGCTGATCAAGCGAT	6671
QY	2100	AGAAATGTTTACAG-----	2113
Db	66717	AGAAATGTTTACAGAGTGAAGGAAATATCTACATTTCAATCAATTTGTGAAACTGT	6677
QY	2114	-----	2128
Db	66777	TGACATGATTAATAGTCTGTGTGCTTGTGATTTCTGTATTAATTAATAGTTTGAATGGGC	6683
QY	2129	GAATGCTGAAGATTTTGGACTGACAGAGACGCTGAAACTGGCAGCTGGTGGTTTATG	2188
Db	66837	GAATGCTGAAGATTTTGGACTGACAGAGACGCTGAAACTGGCAGCTGGTGGTTTATG	6689
QY	2189	ATTATPACACTGTTGAAACTATCTGTGTGACAGGTGTGACAGTCTGACAGATGAAACCCGTG	2248
Db	66897	ATTATPACACTGTTGAAACTATCTGTGTGACAGGTGTGACAGTCTGACAGATGAAACCCGTG	6695
QY	2249	CTCTGTGGGAAGCAACTCTGAGAGAGTCTGTTGTCTATCTGATTTGGTTCAATCCAGAAA	2308
Db	66957	CTCTGTGGGAAGCAACTCTGAGAGAGTCTGTTGTCTATCTGATTTGGTTCAATCCAGAAA	6701
QY	2309	ACAATGCTACTGATGTGAGAACTPACACAAAGATATGAAAGTTTTCTGGTCCAGATCGAG	2368
Db	67017	ACAATGCTACTGATGTGAGAACTPACACAAAGATATGAAAGTTTTCTGGTCCAGATCGAG	6707
QY	2369	GGTGGAATATCACTGAAAGGCTCTGTCTTTGTCATATAA	2406
Db	67077	GGTGGAATATCACTGAAAGGCTCTGTCTTTGTCATATAA	6711

```

DEFINITION      Medicago truncatula clone mth2-155019, WORKING DRAFT SEQUENCE, 4
ACCESSION      AC158210
VERSION        AC158210.11 GI:71061528
KEYWORDS       HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
               Medicago.
REFERENCE      1 (bases 1 to 133779)
AUTHORS        Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
               and Roe,B.A.
TITLE          Medicago truncatula BAC Clone mth2-155019
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 133779)
AUTHORS        Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
               and Roe,B.A.
TITLE          Direct Submision
JOURNAL        Submitted (09-MAR-2005) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE      3 (bases 1 to 133779)
AUTHORS        Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
               and Roe,B.A.
TITLE          Direct Submision
JOURNAL        Submitted (22-JUL-2005) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
COMMENT        ON JUL 22, 2005 this sequence version replaced gi:68342147.
               ----- Genome Center
               Center: Department Of Chemistry And Biochemistry
               The University Of Oklahoma
               Center code:UOKNOR
               -----
               * NOTE: This is a 'working draft' sequence. It currently
               * consists of 4 contigs. Gaps between the contigs
               * are represented as runs of N. The order of the pieces
               * is believed to be correct as given, however the sizes
               * of the gaps between them are based on estimates that have
               * provided by the submittor.
               * This sequence will be replaced
               * by the finished sequence as soon as it is available and
               * the accession number will be preserved.
               *
               * 1 43719: contig of 43719 bp in length
               * 43720 43819: gap of unknown length
               * 43820 83829: contig of 40110 bp in length
               * 83930 84029: gap of unknown length
               * 84030 106272: contig of 22243 bp in length
               * 106273 106372: gap of unknown length
               * 106373 133779: contig of 27407 bp in length.
               *
               Location/Qualifiers
                 source
                   1..133779
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-155019"
                     /clone_lib="Medicago truncatula BAC library H2"
                     43720..43819
                       /estimated_length=unknown
                       83930..84029
                       /estimated_length=unknown
                       106273..106372
                       /estimated_length=unknown
ORIGIN
Query Match 10.1%; Score 243.6; DB 14; Length 133779;
Beet Local Similarity 59.6%; Pred. No. 2,4e-54;
Matches 535; Conservative 0; Mismatches 229; Indels 134; Gaps 2
683 ATTTATTACTGCTATGATGAGTTGTGTGGAGGAAGCTTTGAAGCTTTTACGAGAGGAAG 742
|||||

```

```
Db 132882 ATTGACATTCTGGTGCAATATGCTTCTGATGACCTGATCATGTAGTAGAAGAGG 132941
Qy 743 CAAGTACCTTGACCGGATTTACGTGCAAAATGATGAGCTTTGGAAGAGATCATCTC 802
Db 132942 CAAGCAGCCTTAGACCGGATTTACAAACAAATGATGATGAGCTTTGGAAGATTAACC 133001
Qy 803 CCGCTATGCTTGGAGCTACTTGGCTTACCGCTGGTGTATGATTAACGCTCCGAAAAGC 862
Db 133002 CAGCTTGTATTAAGAACTTTTACCTTGCTTGTATGATGAAACATCGAGCGCGAGAG 133061
Qy 863 TAAATGCTTAAAGCGGTGCGGAATATTTGTGTCTGTGTGGAGAGAGTGACATCAG 922
Db 133062 AAGAAAGCTCCAAAGGTGTCGCAACATTTCTGTGGCAGTTGAGGTGGAGGAGCAG 133121
Qy 923 CTCTTGTGGGGTTTGACCCGAGAAAGTTATGAAATGAGGGCTTTTACGAATGACAG 982
Db 133122 CAATTGCTGGAGTTTACACGAGAGATTCATGAAATGAGGATTCCTGCAATAGAAAG 133181
Qy 983 CTGCTGAGC----- 991
Db 133182 CTGCGCAACAGGTTCATAGAAATATATGTCATGTTTTTATTTTAAATTTCTCCCG 133241
Qy 992 ----- 991
Db 133242 CCCCTTAGTTTTGTATGAAGTTTTTACATTTATCTCTTTCATTAATCTTTTATC 133301
Qy 992 -----AGTTGATCTTTTGTAGTACCCGAGCAATATTCAGCAG 1034
Db 133302 TCCAAATCTTTTTTTTAAAGGTGCAACTTTTGTAGCAACCTTAGCAATATTCAGCTGA 133361
Qy 1035 GTCAATTTGAAGTTTACGAAGTTGCACCTGCTTGTGCTCAAGCTTTTATTTGTAAGA 1094
Db 133362 AAGTTTGAAGCCTTAGGGGTTGCACTTGATGTTGTCAGCAAGCCTTTGTAGTAAGAA 133421
Qy 1095 GCCACACCTTTTACAGATCTGATTAAGCAATTCAGCAACTTCAGAGCCTTAGATAT 1154
Db 133422 GCCACATCTTATCAAGATGCTGATTAATTTTCCATCACTTCAACAACTAAGGTAAAC 133481
Qy 1155 GGGTATGAGATTCCTGCGCATGTTGATGATAC-----ACGAAATATTGGAGATPAG 1208
Db 133482 AATAATGAGAAAGTCCCTCTGTTTACTCCCATGAGAGATGAGAAAGAGAGTTGA 133541
Qy 1209 CTTGCGCTAGAAAGGGAATCTGTGCACTGTTATAGCAAGTATGATGATGATGATGAT 1268
Db 133542 TTTTGCAATTAAGAAAGGGGTGTGTGCACTGTTGTGGGAGCCTTGAATGTGATC 133601
Qy 1269 GTGGTTGGCTTACAGATGAGGATTCACAATATAGGAATCCAGTATTTGAGATTTGT 13328
Db 133602 ATGTTTGGGTCTGATAGTACAGCTCACCTTATCGAAACCCATCTATTTATAGACTTAT 133661
Qy 1329 TTTTGGAATTTCAAAATCGTATGACAAATGATGATCTCCCTGGACTATGCAATTTGGA 1388
Db 133662 TATGGAAGCGCAAAAGGTATAGACAGTGAATCTTCGGAAGCTGTAAATTTGTTGA 133721
Qy 1389 AACCTGTTGGCAGGAGTTGTCTTTCTAGGTTCAAGAGACCAAAAGATTAATAATTT 1446
Db 133722 GACATGTTGATGAGAGTGTCTTTCTAGGTTTAAAGATTAATAAGAGCAAACTTT 133779
```

RESULT 8
AP008208_011/c
WPCOMMENT
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

```
AP008208_009 900001 1010000
AP008208_010 1000001 1100000
AP008208_011 1100001 1210000
AP008208_012 1200001 1310000
AP008208_013 1300001 1410000
AP008208_014 1400001 1510000
AP008208_015 1500001 1610000
AP008208_016 1600001 1710000
AP008208_017 1700001 1810000
AP008208_018 1800001 1910000
AP008208_019 1900001 2010000
AP008208_020 2000001 2110000
AP008208_021 2100001 2210000
AP008208_022 2200001 2310000
AP008208_023 2300001 2410000
AP008208_024 2400001 2510000
AP008208_025 2500001 2610000
AP008208_026 2600001 2710000
AP008208_027 2700001 2810000
AP008208_028 2800001 2910000
AP008208_029 2900001 3010000
AP008208_030 3000001 3110000
AP008208_031 3100001 3210000
AP008208_032 3200001 3310000
AP008208_033 3300001 3410000
AP008208_034 3400001 3510000
AP008208_035 3500001 3610000
AP008208_036 3600001 3710000
AP008208_037 3700001 3810000
AP008208_038 3800001 3910000
AP008208_039 3900001 4010000
AP008208_040 4000001 4110000
AP008208_041 4100001 4210000
AP008208_042 4200001 4310000
AP008208_043 4300001 4410000
AP008208_044 4400001 4510000
AP008208_045 4500001 4610000
AP008208_046 4600001 4710000
AP008208_047 4700001 4810000
AP008208_048 4800001 4910000
AP008208_049 4900001 5010000
AP008208_050 5000001 5110000
AP008208_051 5100001 5210000
AP008208_052 5200001 5310000
AP008208_053 5300001 5410000
AP008208_054 5400001 5510000
AP008208_055 5500001 5610000
AP008208_056 5600001 5710000
AP008208_057 5700001 5810000
AP008208_058 5800001 5910000
AP008208_059 5900001 6010000
AP008208_060 6000001 6110000
AP008208_061 6100001 6210000
AP008208_062 6200001 6310000
AP008208_063 6300001 6410000
AP008208_064 6400001 6510000
AP008208_065 6500001 6610000
AP008208_066 6600001 6710000
AP008208_067 6700001 6810000
AP008208_068 6800001 6910000
AP008208_069 6900001 7010000
AP008208_070 7000001 7110000
AP008208_071 7100001 7210000
AP008208_072 7200001 7310000
AP008208_073 7300001 7410000
AP008208_074 7400001 7510000
AP008208_075 7500001 7610000
AP008208_076 7600001 7710000
AP008208_077 7700001 7810000
AP008208_078 7800001 7910000
AP008208_079 7900001 8010000
AP008208_080 8000001 8110000
AP008208_081 8100001 8210000
```

AP008208_082 8200001 8310000
 AP008208_083 8300001 8410000
 AP008208_084 8400001 8510000
 AP008208_085 8500001 8610000
 AP008208_086 8600001 8710000
 AP008208_087 8700001 8810000
 AP008208_088 8800001 8910000
 AP008208_089 8900001 9010000
 AP008208_090 9000001 9110000
 AP008208_091 9100001 9210000
 AP008208_092 9200001 9310000
 AP008208_093 9300001 9410000
 AP008208_094 9400001 9510000
 AP008208_095 9500001 9610000
 AP008208_096 9600001 9710000
 AP008208_097 9700001 9810000
 AP008208_098 9800001 9910000
 AP008208_099 9900001 10010000
 AP008208_100 10000001 10110000
 AP008208_101 10100001 10210000
 AP008208_102 10200001 10310000
 AP008208_103 10300001 10410000
 AP008208_104 10400001 10510000
 AP008208_105 10500001 10610000
 AP008208_106 10600001 10710000
 AP008208_107 10700001 10810000
 AP008208_108 10800001 10910000
 AP008208_109 10900001 11010000
 AP008208_110 11000001 11110000
 AP008208_111 11100001 11210000
 AP008208_112 11200001 11310000
 AP008208_113 11300001 11410000
 AP008208_114 11400001 11510000
 AP008208_115 11500001 11610000
 AP008208_116 11600001 11710000
 AP008208_117 11700001 11810000
 AP008208_118 11800001 11910000
 AP008208_119 11900001 12010000
 AP008208_120 12000001 12110000
 AP008208_121 12100001 12210000
 AP008208_122 12200001 12310000
 AP008208_123 12300001 12410000
 AP008208_124 12400001 12510000
 AP008208_125 12500001 12610000
 AP008208_126 12600001 12710000
 AP008208_127 12700001 12810000
 AP008208_128 12800001 12910000
 AP008208_129 12900001 13010000
 AP008208_130 13000001 13110000
 AP008208_131 13100001 13210000
 AP008208_132 13200001 13310000
 AP008208_133 13300001 13410000
 AP008208_134 13400001 13510000
 AP008208_135 13500001 13610000
 AP008208_136 13600001 13710000
 AP008208_137 13700001 13810000
 AP008208_138 13800001 13910000
 AP008208_139 13900001 14010000
 AP008208_140 14000001 14110000
 AP008208_141 14100001 14210000
 AP008208_142 14200001 14310000
 AP008208_143 14300001 14410000
 AP008208_144 14400001 14510000
 AP008208_145 14500001 14610000
 AP008208_146 14600001 14710000
 AP008208_147 14700001 14810000
 AP008208_148 14800001 14910000
 AP008208_149 14900001 15010000
 AP008208_150 15000001 15110000
 AP008208_151 15100001 15210000
 AP008208_152 15200001 15310000
 AP008208_153 15300001 15410000
 AP008208_154 15400001 15510000

AP008208_155 15500001 15610000
 AP008208_156 15600001 15710000
 AP008208_157 15700001 15810000
 AP008208_158 15800001 15910000
 AP008208_159 15900001 16010000
 AP008208_160 16000001 16110000
 AP008208_161 16100001 16210000
 AP008208_162 16200001 16310000
 AP008208_163 16300001 16410000
 AP008208_164 16400001 16510000
 AP008208_165 16500001 16610000
 AP008208_166 16600001 16710000
 AP008208_167 16700001 16810000
 AP008208_168 16800001 16910000
 AP008208_169 16900001 17010000
 AP008208_170 17000001 17110000
 AP008208_171 17100001 17210000
 AP008208_172 17200001 17310000
 AP008208_173 17300001 17410000
 AP008208_174 17400001 17510000
 AP008208_175 17500001 17610000
 AP008208_176 17600001 17710000
 AP008208_177 17700001 17810000
 AP008208_178 17800001 17910000
 AP008208_179 17900001 18010000
 AP008208_180 18000001 18110000
 AP008208_181 18100001 18210000
 AP008208_182 18200001 18310000
 AP008208_183 18300001 18410000
 AP008208_184 18400001 18510000
 AP008208_185 18500001 18610000
 AP008208_186 18600001 18710000
 AP008208_187 18700001 18810000
 AP008208_188 18800001 18910000
 AP008208_189 18900001 19010000
 AP008208_190 19000001 19110000
 AP008208_191 19100001 19210000
 AP008208_192 19200001 19310000
 AP008208_193 19300001 19410000
 AP008208_194 19400001 19510000
 AP008208_195 19500001 19610000
 AP008208_196 19600001 19710000

Query Match 7.5%; Score 180.8; DB 15; Length 110000;
 Best Local Similarity 55.1%; Pred. No. 3.1e-37;
 Matches 510; Conservative 0; Mismatches 292; Indels 124; Gaps 3;

QY 731 AGGAGGAAGCAAGTACCTTTGACCGAATTTAGTCACAAATTGATGAGACTTTGG 790
 DB 81260 AGGAAGTGAAGCAAGCAATCTCGACCTGATCTGTTTCACAGATTGAACTCTCG 81201
 QY 791 AAGAGATCACTCCGCGTTATGTTTGGAGCTTACTTGGCTTGTGATGATTACG 850
 DB 81200 AGGAGTTACACTGCTGTGATTGAGACTTCTCCCTTCATTGACACAGACATC 81141
 QY 851 CTGCGAAAAGACTTAATGTTTAAGCGGTGTCGGAATTTTGTGTCGTCTGTTGAGAG 910
 DB 81140 ATTAAGAAAGCGCAAGAAAGGCTTCAAGTGCAGAAACATTTTGTGAGCGTTGGCAGAG 81081
 QY 911 GTGAGCATCAGCTCTTGTGTGGGGGTTTGACCCGTGAGAAATTATGATGAGCGCTTT 970
 DB 81080 GAGGATTGCTACCTGTGAGAGAGATTCTTCGTGTAACCTTCATGAACGAGCTTTT 81021
 QY 971 TACGAATGACAGCTGCTGAGAGTT----- 996
 DB 81020 TGAGATGACATCAATTGAACAGGTTCTGACATGTTTGTATTGAGACAGCTGCTAT 80961
 QY 997 ----- 996
 DB 80960 TTATATTGACAGAGACATCTTTGAAMACATGTTAATCCCTTCTGTTACCTCCAG 80901
 QY 997 ---GATCTTTTGTAGTACCCCAAGCAATATTCGACGAGTGCATTTGAAGTTACGAA 1053

Db 80900 ATGATTTCTTTCAAAAACACGGAATAGACTTCCTGTAATGGTTGTAATTTACAT 80841
 Oy 1054 GTTGACATTTGCTCTTGTGGCTCAAGCTTTATTTGTAAGAACCCACCTTTTACAGAT 1113
 Db 80840 GTAGACATTTGACATGTCCTCCTCAGCAATTTATGTAAGAAAAGCCACAAATTCATGATG 80781
 Oy 1114 GGTGATTAAGCAATTCACGACATCTGACGAGGCTAAGGTATAGGCTATGAGATTCCTGCG 1173
 Db 80780 GCGGATGATCTTTTGAACACTCCAGAAATTTCA-----C 80745
 Oy 1174 ATGTTGATGATACACGGAATATTTGGAGATAGACTTGCTGCTAGAAAAGGCACTGT 1233
 Db 80744 ATAGGTTCCATTATGCTTATGATATGATGAGACTTGCAATGGAAGGCAATTCCTGC 80685
 Oy 1234 GCATGCTTATAGGCAAGTTGATGAATGCCATATGCTGTTGGGCTTAAACAGTGAAT 1293
 Db 80684 TCATTGCTATGTCGGAGATGTTGACAGATCAGAAATGTGCTTGGAAATTAATATGATGCT 80625
 Oy 1294 TCACAATTAAGGAATCAGACTATTTGGAGTTTGTGTTGAGAAATCAATC---GTGAT 1350
 Db 80624 TCACCAATACAGAACCCCAAAATTTAGAGTTTATTTGACCAACTTACATCACTGA 80565
 Oy 1351 GACAAATGATATCTCCCTGCACTATGCAAAATTTGTTGAAAACCTGTGGCAGGCTGTGC 1410
 Db 80564 GAGATGATCTTCTTCCAGGGCTGTCCAAAGCTTTGAGAGCTTGCTTATCTTTGAGGTT 80505
 Oy 1411 TTTCTAGGTTGAGAGACACCAAGATTAATAAATCTGGGGAATCTATGATGAT 1470
 Db 80504 TTTCTAGGAGACAGATATCTCGGGGCACTGCACTTGAAGATTAATCAATTAATCAAT 80445
 Oy 1471 CCTATGTTTGTGATTAATTTGAAAAGAGTAGGAGTTCAGGGTCTCTCTTATGCTGCT 1530
 Db 80444 CCAAGAGTTTAAAGTAACTTAAAGAAAGATGAGAGGCTGCTCTTCTCATTTGGCTGCT 80385
 Oy 1531 GCTGCAACTATGGCAAGATTTGAGC 1556
 Db 80384 GCTGCTGCTATGCAAAACTTGCTGC 80359

RESULT 9
 AP004885/c 150462 bp DNA linear PLN 15-SEP-2004
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 DEFINITION PAC clone: P0575F10.
 ACCESSION AP004885
 VERSION AP004885.3 GI:41053009
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
 clone: P0575F10
 JOURNAL Published Only in Database (2002)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submision
 JOURNAL Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 On Jan 21, 2004 this sequence version replaced gi:38142429.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mtc.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), Glimmer
 (http://www.tigr.org/tigr/tdb/glimmer/glimr_form.html), RiceHMM
 (http://rpg.dna.affrc.go.jp/RiceHMM/), SplinePredictor
 (http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), sim4

FEATURES

source

(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI Nonredundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DBJ. Protein homologues of the coding
 regions were searched against NCBI Nonredundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from T7 to SP6 of the PAC clone.
 This sequence of P0575F10 clone has an overlap with OJ1020_C02
 (DBJ: AP004078) clone at 5' end and with P0482P12 (DBJ: AP005311)
 clone at 3' end. Detailed information on overlap and assembly
 quality together with annotation of this entry is available at
 http://rpg.dna.affrc.go.jp/Genomeseq.html.

Location/Qualifiers

1. 150462
 /organism="Oryza sativa (japonica cultivar-group)"
 /molecule="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0575F10"
 /complement(4987..8112)
 /gene="P0575F10.1"
 /complement(join(4987..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8112))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK106553"
 /complement(join(5003..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8084))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK073059"
 /complement(join(5315..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..7876))
 /gene="P0575F10.1"
 /note="contains EST(s): AU070160(B51534)
 contains full-length cDNA(s): AK106553, AK073059"
 /codon_start=1
 /product="putative eukaryotic translation initiation
 factor 3, subunit 1 alpha, 35kDa"
 /protein_id="BAD07941.1"
 /db_xref="GI:41053010"
 /translation="MEDWSDPFAVPSAKAPLPSKNAEDVEDVDYKSWEEER
 EKRPPEVETAPKPSGKAQKQKQSEVVDLDPALKEKLEKQELVEAD
 EKSTTRPKGDKGSKSLDPTPKESDPAEYALILANGLRVEKSFHYMGLIKVMR
 LSWASIKGDAKDQISSSTAIINKEIKAKKEAAAGKKQKQKQHIENKDDPTFG
 KKNFDDPDERYDFM"
 /complement(10186..13754)
 /gene="P0575F10.2"
 /complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="start and end point are not identified"
 /complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="contains EST(s): AU058423(B51136), AU189368(B51136)"
 /codon_start=1
 /product="plastid division protein precursor"
 /protein_id="BAD07942.1"

CDS

gene
mRNA

```

/db_xref="GI:41053011"
/translation="MEGFHNLARPNGAPPAFSLPRPRPRRRPRPPPPAGRAASR
MARIADPDLILPTAASDPSPAPAPAAAPSAAPRPLPPDAERSLPQNDPKYKL
GAEPHFLGDIIRAFERIKAPPOYGTSTDLVGRMLDIADTLNQNSTRQYDRA
LSNRREALTMDIAMDKEGELAVLTGBQLIDRPKPKODVLAAMLAAYDLSR
DAAASBPVDVIGCEVLERALKLOEBDGNALPDLISQIDETLEETPCVLESLT
PITENHKKRQEGLOGARNTLWVGSGFSAFBAFMAFLMMSICOMDF
SKTPNSIPPEPFIYVVALAHVAQAIISKRPOFLMADLFEOLQKFNIGSHAYDNE
MDLALERAPSLVGDVSKRMVLGIDNESPYPDPKYLEPIYNSISERNLRL
CKLEWTLIFEVPRSDTRKMQPRLDYDDDEBVTSLYLRMGEGASHLAALAK
LGAQATVLTGVSNALQAFNKVPELLEQDRAMENYKQPGYLENFOENAPAD
SRNALKISAGALFALAVIYAKYLERKPLSARSHSVAVANSVSDTPALDE
DPVHIDMDAKLAEIVRKQSIKSKALGPEHSAVLOEVLDMKLVKMDRAETIER
HGMFWETLSVDVITISLIDGRRAVETIDBAGQLTVTEPRNNDSYDTYTRY
EMAFKXGKTEGAVLKS"
complement(16401..17341)
/gene="P0575F10.3"
/complement(join(16401..16739,16841..>17341))
/gene="P0575F10.3"
/notes="start and end point are not identified"
complement(join(16401..16739,16841..17341))
/gene="P0575F10.3"
/notes="similar to Arabidopsis thaliana chromosome 3,
Ac3g61870"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07943.1"
/db_xref="GI:41053012"
/translation="MMAAALPAVAHAIITCSASNKSPSPAPQOQTTTATGSPAA
LPLRLTTAAATAAALAPALAPALAAAGGFLIEGSAVLPLWGGIATLWAG
YLGOMRWRTTIDEBINELKPKAAPAAAPPAVAAAGSSSSSSPPAPSPHPI
KIDELTEBERKLIKSGSPDRHPNAGSLIGLGTYESVCGALNPPRTGKLPHPPI
GAATITVMAAALVPMQKNETASLHIALNAINVLTWQIPGLVGRVFEFT
TF"
complement(18499..18942)
/gene="P0575F10.4"
/complement(<18499..>18942)
/gene="P0575F10.4"
/notes="start and end point are not identified"
complement(18499..18942)
/gene="P0575F10.4"
/notes="contains EST(s) : AU184851(R10374),AU184850(R10374)"
/codon_start=1
/product="putative BF-hand Ca2+-binding protein CCD1"
/protein_id="BAD07944.1"
/db_xref="GI:41053013"
/translation="NMHHQIIRSTTAAEQEASAGGGGGEYEDLMPVAGRLGAEGL
LSLRAGFRLLADPARGAITAESLRSASVLGLGGGGEMTVEBAAMVREBDQGD
GALSEAFCTIWRSLSPGIMGDAEWLBEALDELRLSPPPPA"
21937..23552
/gene="P0575F10.5"
/join(<21937..22225,22326..22648,22736..23049,
23393..23552)
/gene="P0575F10.5"
/notes="start and end point are not identified"
join(21937..22225,22326..22648,22736..23049,23393..23552)
/gene="P0575F10.5"
/notes="similar to Arabidopsis thaliana chromosome 5,
Ac5g24330"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07945.1"
/db_xref="GI:41053014"
/translation="MGPATPPLRRRTARPARATBAGSGGDDDDVRCACSGESA
ABLLCDGCDRGHIFCLRPILPRVAGWFCSPSCSPSHSKSKAAKQFPLVO
TKIVDFKIQRGRAALAAABESSSEGRKRVKRGALRVSKRKLPPSPDPAR
LRQMASLATAATGAVSNELTYVPGMPPRANRALSAGMUVPKEDVETNLCK
RMAAGSEWPLVVDVPGVGTVEADRFKDLITTEYVYGVYDVLRRRHDDSMRT
LLSATPSSSLVCPDKRSIARFINGINNHTPDGRKKQNLCKVRDVGECVLLVA
NNDISGERLVYDNGSERHYPTHHFV"
complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/gene="P0575F10.6-1"

```

```

mRNA      complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/gene="P0575F10.6-1"
/notes="supported by full-length cDNA(s) : AK121372"
complement(join(24491..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28147))
/gene="P0575F10.6-1"
/notes="supported by full-length cDNA(s) : AK062255"
complement(24563..28147)
/gene="P0575F10.6-2"
/complement(24563..28147)
/gene="P0575F10.6-2"
/notes="contains full-length cDNA(s) : AK065148
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
complement(join(24787..24817,24913..25266,25398..25633,
25720..25832,27054..27159))
/gene="P0575F10.6-1"
/notes="contains EST(s) : AU031017(B60571),AU093466(B60571)
contains full-length cDNA(s) : AK121372,AK062255"
/codon_start=1
/product="putative arginine/serine-rich splicing factor

Query Match      7.5%; Score 180.8; DB 15; Length 150462;
Best Local Similarity 55.1%; Pred. No. 3.2e-37;
Matches 510; Conservative 0; Mismatches 292; Indels 124; Gaps 3;

mRNA      731 AGGAGAAAGGACAGTACGCTTGACCGGATTTAGTCGACAAATTGATGACCTTTGG 790
12351 AGGAATATGAGGACAACTCTCGACCTGATCTGCTTACAGATGAGAACTCTCG 12292
791 AATAGATCATCTCCGCTTATGCTTGTGAGCTACTTGGCTTACCGCTTGGATGATTACG 850
12291 AGGAGATTACACCTCGCTGTGATGAGCTTCTCCCTTCTTATGACACAGACATC 12232
851 CTGCAAAAAGACTAAATGCTTTAAGCGGTGCGCAATATTTGTGCTGTGAGAG 910
12231 ATTAGAAGCGCCAAAGAGGCTTCAAGTCCGAGAAACATTTGTGAGAGGTTGGCAGAG 12172
911 GTGAGACATCAGCTCTTGTGGGGTTTGAACCGTGAAGAGTTTATGATGATGAGCGCTTTT 970
12171 GAGGTATTGCTACGTTTGGAGGAGATTTCTGTAAGCCTTCATGAAAGAGGCTTTT 12112
971 TAGATGACAGTCTGACAGGTT----- 996
12111 TGAGATGACATCAATTAAGACAGTTCGTACATGTTGTTTATGAGCAGCTGCTTAT 12052
997 ----- 996
12051 TTATATTGACAGACATCTTTGAAAACATGTAATCCCTTCTGTACACTCCAG 11992
997 ---GATCTTTTGTAGCTACCCGACCAATATTCAGACAGTATTTGAAGTTTACGA 1053
11991 ATGATTTCTTTTCAAAAACCGAATGACATTCCTCGTAATGTTGAAATTTACAAT 11932
1054 GTTGCACTTGCTTGTGTGCTCAAGCTTTATTTGTAAGAACCAACCTTTTACAGAT 1113
11931 GTAGCACTTGCAATGCTGCTCAAGCAATTAATGTAAGCAAGCAATTCATCATGATG 11872
1114 GCTGATAGCAATTCAGCAACCTTGACAGGCTTAGGTATGAGATTGCTGCG 1173
11871 GCGGATGATCTTTTGAACAATCCGAAAGTTCA-----C 11836
1174 ATGTTGATGATACAGGAATATTGGAGATAGACTTCGCTTGAAGAGGAGCTCTGT 1233
11835 ATAGTTTCATATATCTTATATGATATGAGACTTGCATTTGAAAGGGGATCTCAG 11776
1234 GCACTGCTTATAGGCAAAATTGATGAATCCGATATGTTGGCTTAAAGCTGAGAT 1293
11775 TCATTGCTATGTCGAGATGTGAGCAAGTCAAGATGTGCTTGAATTAATATGATCT 11716
1294 TCACATATAGGAATCAGCTATTTGAGATTGTTTGGAGATTCATATC---GTGAT 1350

```

Db 11715 TCACCATACAGAGACCCCAAAATTCCTAGATTATGTGACCAACTCTACATCACTGAA 11656

Qy 1351 GACATGATGATCTCCCTGGACTATGCAATTTGTGGAAACCTGTTGGAGGGGTTC 1410

Db 11655 GAGAAATGATCTTTCTCCAGGGCTGTGACCTTTGGACCTTGCTTATCTTTAGAGTT 11596

Qy 1411 TTTCCTAGGTTTCAGAGACCAAGATTAATAAATTTAACTCGGGGACTATATGATGAT 1470

Db 11595 TTTCCTAGAGAGCAGAGATCTCGGGGCACTGACATTCAGACTTGGAGATTCTACGATGAT 11536

Qy 1471 CCTAGGTTTGAATTAATCTGAAAGATGAGAGTACTGAGGTTCTCTTACTGCTT 1530

Db 11535 CCAGAAAGTTTAACTGCTAGAAAGAGATGAGAGGTGTGTGCTCTTCACTTGGCTGCT 11476

Qy 1531 GCTGCAACTATGGCAAGATTGAGC 1556

Db 11475 GCTGCTGCTATGCAAAACTGGTGC 11450

RESULT 10

LOCUS DQ022571 550 bp DNA linear STS 28-MAY-2005

DEFINITION Beta vulgaris chromosome 1, sequence tagged site.

ACCESSION DQ022571

VERSION DQ022571.1 GI:66394762

KEYWORDS STS.

SOURCE Beta vulgaris

ORGANISM Beta vulgaris

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Caryophyllales; Amaranthaceae; Beta.

1 (bases 1 to 550)

Friesen, T.L., Welland, J.J., Aasheim, M.L., Hunger, S., Borchardt, D.C. and Lewellen, R.T.

Identification of a SCAR marker associated with Bm the Beet mosaic virus resistance gene on chromosome 1

Unpublished

2 (bases 1 to 550)

Welland, J.J. and Friesen, T.L.

Direct Submission

Submitted (29-APR-2005) Sugarbeet and Potato Research, USDA-Agricultural Research Service, 1307 18th St. N, Fargo, ND 58105, USA

FEATURES

source location/Qualifiers

1..550

/organism="Beta vulgaris"

/mol_type="genomic DNA"

/cultiivar="C719"

/db_xref="taxon:161934"

/chromosome="1"

/note="fwd primer seq: GACTGAGCTGTAAGCAGCTGT; rev primer seq: GGAAGCATTTCACTCTTTATGCT; fwd primer name: Rbm05fwd; rev primer name: Rbm05rev"

1..550

/note="SCAR marker associated with resistance to Beet mosaic virus"

ORIGIN

Query Match 6.9%; Score 165.4; DB 10; Length 550;

Best Local Similarity 74.6%; Pred. No. 2.5e-33;

Matches 208; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 718 TTGAAGCTTTAAGAGAGAGCAAGTACCTGACCGATTTAAGTGCACAAT 777

Db 102 TTGACTATTTTGAAGAGAGGAGGTCAAGTAGTTCGCCGATTTAAGAGCAGAGGTT 161

Qy 778 GATGAGCTTTGAGAGATCACTCCGCTTATGCTTGAAGCTACTTGCTTACCGCTT 837

Db 162 GAGAGAGCATTTGAGAGATCAAGACCCCTTATGCTTGAAGCTTTGGCTTGGCCCTC 221

Qy 838 GGTGATGATTACGCTGCGAAAGACATTAATGCTTTAAAGCGGTGCGGAAATTTTGG 897

Db 222 AGTGATGAATTTGGAAGAAAGAGATGAAGGCTTCAAGGTGACCAATATATATYMG 281

Qy 898 TCTGTTGAGAGAGTGGACATCACTCTGTGGGGGTTTGAACCCGTGAGAGTTTATG 957

Db 282 TCTGTTGAGAGAGTGGTCCCGAGCTATAGCTGGGTGTTTACTGCTGAGATTTGATG 341

Qy 958 AATGAGGCTTTTACGAATGACAGCTGCTGAGCAGGTT 996

Db 342 AATGAGGCTTTTTCATATGACAGACGCTGAGCAGGTT 380

RESULT 11

LOCUS AC155597 184752 bp DNA linear HTG 25-JAN-2005

DEFINITION Zea mays strain B73 clone ZMMB8c0216X08, *** SEQUENCING IN PROGRESS

ACCESSION AC155597

VERSION AC155597.2 GI:58082456

KEYWORDS HTG; HTGS_PHASE1.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogonaceae; Zea.

1 (bases 1 to 184752)

Chan, A.P., Perrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecherback, T.R., Feldbljum, T.V., Rabinowicz, P., Frazer, C.M., Schubert, K., Sanmiguell, P., Ma, J., Pontaroli, A.C., Rohlfing, T., Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J. and Quackenbush, J.

Direct Submission

Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850

3 (bases 1 to 184752)

Chan, A.P., Perrea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecherback, T.R., Feldbljum, T.V., Rabinowicz, P., Frazer, C.M., Schubert, K., Sanmiguell, P., Ma, J., Pontaroli, A.C., Rohlfing, T., Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J. and Quackenbush, J.

Direct Submission

Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850

On Jan 25, 2005 this sequence version replaced gi:57863118.

----- Trace submission

Center name: TIGR

Seq_id: ZGCO

Web site: <http://www.tigr.org/cdb/cgi/maize/>

Contact: maize@tigr.org

----- Project information

NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2210: contig of 2210 bp in length

1 2211 2310: gap of unknown length

1 2311 9941: contig of 7631 bp in length

1 9942 10041: gap of unknown length

1 10042 16745: contig of 6704 bp in length

1 16746 16845: gap of unknown length

1 16846 32394: contig of 15549 bp in length

1 32395 32494: gap of unknown length

*	33495	35494: contig of 3000 bp in length
*	33495	35594: gap of unknown length
*	33595	40051: contig of 4457 bp in length
*	40052	40151: gap of unknown length
*	40152	41379: contig of 1228 bp in length
*	41380	41479: gap of unknown length
*	44180	44517: contig of 3038 bp in length
*	44518	44617: gap of unknown length
*	44618	45564: contig of 947 bp in length
*	45565	45664: gap of unknown length
*	49441	49440: contig of 3776 bp in length
*	49441	49540: gap of unknown length
*	49541	52698: contig of 3158 bp in length
*	52699	52798: gap of unknown length
*	53799	55964: contig of 3166 bp in length
*	55965	56064: gap of unknown length
*	55965	59236: contig of 3172 bp in length
*	59237	59336: gap of unknown length
*	59337	70052: contig of 10716 bp in length
*	70053	70152: gap of unknown length
*	70153	84603: contig of 14451 bp in length
*	84604	84703: gap of unknown length
*	84704	86418: contig of 1715 bp in length
*	86419	86518: gap of unknown length
*	86519	88069: contig of 1551 bp in length
*	88070	88169: gap of unknown length
*	88170	89661: contig of 1492 bp in length
*	89662	89761: gap of unknown length
*	89762	91357: contig of 1596 bp in length
*	91358	91457: gap of unknown length
*	91458	132727: contig of 41270 bp in length
*	132728	132827: gap of unknown length
*	132828	137468: contig of 4641 bp in length
*	137469	137568: gap of unknown length
*	137569	163943: contig of 26275 bp in length
*	163944	163943: gap of unknown length
*	163944	169588: contig of 5915 bp in length
*	169589	169598: gap of unknown length
*	169599	176100: contig of 6142 bp in length
*	176101	176200: gap of unknown length
*	176201	183079: contig of 6879 bp in length
*	183080	183179: gap of unknown length
*	183180	184752: contig of 1573 bp in length.
FEATURES		
source		
1..184752		
/organism="Zea mays"		
/mol_type="genomic DNA"		
/strain="B73"		
/db_xref="taxon:4577"		
/clone="ZMMBC0216K08"		
2211..2310		
gap	estimated length=unknown	9942..10041
gap	estimated length=unknown	16746..16845
gap	estimated length=unknown	32395..32494
gap	estimated length=unknown	35495..35594
gap	estimated length=unknown	40052..40151
gap	estimated length=unknown	41380..41479
gap	estimated length=unknown	44518..44617
gap	estimated length=unknown	45565..45664
gap	estimated length=unknown	49441..49540
gap	estimated length=unknown	52699..52798
gap	estimated length=unknown	55965..56064

Query Match	Best Local Similarity	Score	DB	Length
gap	/estimated_length=unknown	59237.59336		
gap	/estimated_length=unknown	70053.70152		
gap	/estimated_length=unknown	84604.84703		
gap	/estimated_length=unknown	86419.86518		
gap	/estimated_length=unknown	88070.88169		
gap	/estimated_length=unknown	89662.89761		
gap	/estimated_length=unknown	91358.91457		
gap	/estimated_length=unknown	132728.132827		
gap	/estimated_length=unknown	137469.137568		
gap	/estimated_length=unknown	163844.163943		
gap	/estimated_length=unknown	169859.169958		
gap	/estimated_length=unknown	176101.176200		
gap	/estimated_length=unknown	183080.183179		
gap	/estimated_length=unknown			
ORIGIN				
Query Match	6.8%; Score 163.4; DB 14; Length 184752;			
Best Local Similarity	54.0%; Pred. No. 1.8e-32; Indels 135; Gaps 3;			
Matches 506; Conservative 0; Mismatches 296;				
721	AGAGAGAAAGAGACCAAGTAGCCTTGCACCGGATTTACGTGCACAAATTGATGAGACTTTGG	750		
108936	AGAGAGATGTGTCAAGCAATCTTGCACCTGAACCTGTTGCACAGATTGATGAACCTTTGG	108999		
791	AAGAGATCACTCCGGCTTAATGTCTTGGAGCTACCTTGGCTTACCGCTTGTGTGATGATTAACG	850		
108996	AGGAGATTACACCTCGTGTGTATGTGAGACTTCTTGCTCTTCTCTACGTAGAAAACATA	109055		
851	CTGCGAAAAGACTAAATGTTTAAACGGGTGTGCGAAATTTTGTGTCTGTGTGAGAGAG	910		
109056	AAATATTAACGCCAAGAGGTCTGCAAGGTGCAAAAACATATTTGTGAGTGTGCGACGGG	109115		
911	GTGAGACATCAAGCTGTGTGTGGGGGTTGACCCGTGAGAACTTATGAATGAGCGCTTT	970		
109116	GTGTATTTGTCTACTGTGTGAGAGAAATTTTCTGTGAGGCTTTCATGAATGAGCTTTCT	109175		
971	TACGAATGACAGCTCTGAGCAGTTGATCTTTTGTGACT	1011		
109176	TGCAGATGACATCACTGAGCAGTTGTGTCTACCAAGTTTAATGAGCAACTACTGT	109235		
1012		1011		
109236	GTGTGTAGTCTGTCCATTTTACTTGAACACAGTTTTTAAAGATCTTCACTCATTTTCTGT	109285		
1012	-----ACCCCAAGCAATTTCCAGCAGACTATTG	1042		
109296	TTATATCTCAGATGATTTCTTCTCTAAACACCGAATGACATCACCTGATGATGTTTG	109355		
1043	AAGTTTACGAAGTTGCACTTGTCTTGTGGCTCAAGCTTTTATTTGTTAAGAACCCACAC	1102		
109356	AGATTTAATAGTGTGCCACTTGGCCACATTTGCTCAAGCAATTCGAAGTAAAGGCCACAA	109415		
1103	TTTATCAGAGATGCTGATAGCAATTTCCAGCACTTTCAGCAGCTTAAGGTAATGCTATGG	1162		
109416	TCATCATGATGGCAGATGAACCTTTCGAACCACTTCAGAAAGTTCAA	109461		
1163	AGATTCTGCGATGTTGTATGATACAGGAATTAATTTGGAGATTAAGCTTGTGTATGA	1222		
109462	-----TATAGTCTCTCATATCTCTTAAGAAAACGAGATGACCTTCGATTAAGAAA	109511		

```

QY      1223 GGGAGCTCTGTGACCTGCTTATAGGCAAGTGTGATGATCCGTATGTGGCTTAG 1282
          |||||
Db      109512 GGGACACTCTGCTGTTACTTGTGGGGGATATTGCAATTCAGAAATGTGGCTTGAATTG 109571
          |||||
QY      1283 ACGTGTGAGATTTCACATATATAGGAATTCAGCTATTTGGAGTTTGTGGAAATTCAA 1342
          |||||
Db      109572 ATATAGTGTCTCTGTCTATATAGGACCCCAAAATATCGATTGTGTGTAATTAACCTTA 109631
          |||||
QY      1343 ATGTG---GATGCAATGATGATCTCCCTGAGCTATGCAATTTGTTGAAACCTGGTTGG 1399
          |||||
Db      109632 GCATGAAAGAAAGAAATGATCTTCTTCAGGGCTGTGCAAGCTCTTGGAAACATGGCTTG 109691
          |||||
QY      1400 CAGGGGTGTCTTCTTCTAGGTTCAAGACACAAAGATATAAAATTTAACTCGGGGACT 1459
          |||||
Db      109692 TCTCCGAGGTATTCCTCCAGAGAGAGATATCCGAGGAGACGAGTTCACTGGAGACT 109751
          |||||
QY      1460 ACTATGATGATCTCTATGCTTTTGAAGTTTCTTGGAAAGTGGAGATGTTCAAGGTTCTC 1519
          |||||
Db      109752 ATTACGACGACCCAAAGTTCTTAAGCTTATAGAAAGATGGAAGCGGTGTGCTTCCC 109811
          |||||
QY      1520 CTTTATGCTGCTGCTGCAACTATGAGCAAGATTTGAGC 1556
          |||||
Db      109812 ATTATGCTGACGCTGCCCTATAGCAAAACTTGTGC 109848
          |||||

RESULT 12
AC160013 117818 bp DNA linear HTG 02-JUN-2005
LOCUS Medicago truncatula clone mth2-162b23, WORKING DRAFT SEQUENCE, 4
DEFINITION
ordered pieces.
AC160013
VERSION AC160013.10 GI:66865008
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
            Burkariya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE 1 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-162b23
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2005) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE 3 (bases 1 to 117818)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2005) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT On Jun 2, 2005 this sequence version replaced gi:66841540.
            ----- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

```

```

FEATURES
source
1..117818
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/CD_xref="taxon:3860"
/clone="mth2-162b23"
/clone_1fb="Medicago truncatula BAC library H2"
1816..18915
/estimated_length=unknown
21098..21197
/estimated_length=unknown
35202..35301
/estimated_length=unknown

* 1 18815: contig of 18815 bp in length
* 18816 18915: gap of unknown length
* 18916 21097: contig of 2102 bp in length
* 21098 21197: gap of unknown length
* 21198 35201: contig of 14004 bp in length
* 35202 35301: gap of unknown length
* 35302 117818: contig of 82517 bp in length.
Location/Qualifiers

ORIGIN
Query Match 6.0%; Score 144.2; DB 14; Length 117818;
Best Local Similarity 68.3%; Pred. No. 2.9e-27;
Matches 200; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY      2114 AGGTTTGGATGGGCGCATGCTGAAGATTGGAGCTGACAGACAGCTGAACCTGGCGACG 2173
          |||||
Db      1085 AGGTGTGGATCGGCGCAATCTTTGAAATATGACATGATGACAGCTGATGAGATTTGACAGAC 1144
          |||||
QY      2174 TTGGGTTGGTTATGATATATATACATCTTTGAAACTATCTGTTGACAGTGTGACGTCTCG 2233
          |||||
Db      1145 TTGGTTGTCATATATGATGATCACTCACTTGAAGATCTCAACATGACAGTGTGACCATATTCAC 1204
          |||||
QY      2234 CAGATGAAACCCGTCTCTGTGTGGAAGCACTCTGGAAGAGTCTGTTCTATCTGATTT 2293
          |||||
Db      1205 AGAATGGGGGGGGGTGAGTAGTAGGAAACCACTTCAAAAGTCTACCACTTCACTGCTG 1264
          |||||
QY      2294 TGGTTCATCAGAAACAAATGCTACTGATGTGAGAACTACACAACTGATGAGAAATTT 2353
          |||||
Db      1265 TTGTCATTCACAGCATGCTACTTCCACAGCAGAACTTCAACACAGATATATGAATTT 1324
          |||||
QY      2354 TCTGTCCAGTCAAGGTGAGGTAATTCATGAAAGCTGTGTTTCATCATTA 2406
          |||||
Db      1325 CTTTTCAGATTCAGGGGTGAAATATATGAAGAGCTGTCTTGTAGTGTTAA 1377
          |||||

RESULT 13
AC094893/c 221756 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-6B10, *** SEQUENCING IN PROGRESS ***
DEFINITION
AC094893
VERSION AC094893.5 GI:30466744
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 221756)
AUTHORS Muzny,D.,Marie., Metzker,M.,Lee., Adamson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Behnmed,F.,
            Biswalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Bunay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Cessari,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Day-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Becotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```



```
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center Project name: GBOM
Center clone name: CH230-6B10
----- Summary Statistics
Assembly program: ATras;
```

```
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 221756: contig of 221756 bp in length.
```

Location/Qualifiers
1. .221756

```
/note="wgs_end_extension
```

```
/note="clone_boundary
```

```
end_sequence: BH3575588"
182948. .183789
/note="clone_boundary
clone_end: Sp6
site: EcoRI
```

end_sequence: BH357591"

2.6%; Score 62; DB 14; Length 221756;
43.3%; Pred. No. 7.9e-05;

conservative 0; Mismatches 380; Indels 0; Gaps 0;

503

TGCTGCTGTGTTGTTGCTGCTGCTATGCTGTTGTTGTTGC 207932

GGGCTCTCTGTGTATTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGGTTGG 569

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

bioRxiv preprint doi: <https://doi.org/10.1101/207872>; this version posted November 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TTCTGCTTAAGGAGAGGTTGCCCTAAGTCGTTTAAGCAAGATGTGGTTTAACTAT 629

21GCTGTTGCTNGTTGTTGTTGCTNGCTGCTGTTGCTNGTNGTTGTTGT 207812

ATGCGTTTCTCGATGTCCTCGAGGATGCTATGGCATTGGATCCACCTGATTTTAT 689

The image displays a gel electrophoresis result with 12 lanes. Lanes 1, 2, and 3 contain molecular weight markers, showing a series of horizontal bands at different positions. Lanes 4, 5, and 6 show PCR products for different samples, with bands appearing at various heights. Lanes 7 through 12 show sequencing results, with bands appearing at specific positions corresponding to the DNA sequence being analyzed.

TGCTGTGTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGT 207752

TTATGAGTTTGTGAGGAAGCTTTTACAGGAGGAAGCACTAG 749

[illegible]

TTGTGCTGCTGCTGTTGTTGCTGCTGCTGCTGTTGTTGTTGCTGT 207692

AACCGGATTTCAGTCACCAATTGATGAGACCTTTGGAAGAGATCACCTCCGCGTTA 809

_____ 2
_____ 3
_____ 4
_____ 5
_____ 6
_____ 7
_____ 8
_____ 9
_____ 10
_____ 11
_____ 12
_____ 13
_____ 14
_____ 15
_____ 16
_____ 17
_____ 18
_____ 19
_____ 20
_____ 21
_____ 22
_____ 23
_____ 24
_____ 25
_____ 26
_____ 27
_____ 28
_____ 29
_____ 30
_____ 31
_____ 32
_____ 33
_____ 34
_____ 35
_____ 36
_____ 37
_____ 38
_____ 39
_____ 40
_____ 41
_____ 42
_____ 43
_____ 44
_____ 45
_____ 46
_____ 47
_____ 48
_____ 49
_____ 50
_____ 51
_____ 52
_____ 53
_____ 54
_____ 55
_____ 56
_____ 57
_____ 58
_____ 59
_____ 60
_____ 61
_____ 62
_____ 63
_____ 64
_____ 65
_____ 66
_____ 67
_____ 68
_____ 69
_____ 70
_____ 71
_____ 72
_____ 73
_____ 74
_____ 75
_____ 76
_____ 77
_____ 78
_____ 79
_____ 80
_____ 81
_____ 82
_____ 83
_____ 84
_____ 85
_____ 86
_____ 87
_____ 88
_____ 89
_____ 90
_____ 91
_____ 92
_____ 93
_____ 94
_____ 95
_____ 96
_____ 97
_____ 98
_____ 99
_____ 100

ATGCTGCTGTTGTTGCTGCTGCTGTTGTTGTTGCTGTTGT 207632

GGAGCTACTTGCTTACCGCTTGGTGATGATTACCGCTGCGAAGAAGACTTAATGG 869

100

TTGCTGCTGTTGTTGCTGCTGCTGCTGCTGTTGTTGTTGCTGCTGC 207572

ATGTTTGGTCTCTGTTGGAGGAGGTTGAGCATTCAGCTTTGTT 929

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Steinle, M., Strong, R., Sutton, A., Swack, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanai, K., Valera, R., Vera, V., Villana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, Y., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 232974)
Morley, K.C.
Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 232974)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:24942402.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLOG
Center clone name: CH230-22EB13
----- Summary Statistics
Assembly program: Atlas 3.0/
Consensus quality: 198057 bases at least Q40
Consensus quality: 202743 bases at least Q30
Consensus quality: 205785 bases at least Q20
Estimated insert size: 207912; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 227401: contig of 227401 bp in length
* 227402: gap of unknown length
* 227502: contig of 1272 bp in length
* 228774: gap of unknown length
* 228874: gap of unknown length
* 231053: contig of 2180 bp in length
* 231054: gap of unknown length
* 231153: contig of 1821 bp in length.
* 231154
Location/Qualifiers
1 . 232974
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

FEATURES
source

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 21, 2006, 13:05:25 ; Search time 7918 Seconds
(without alignments)
4733.071 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGLSPFLCRLP.....YEVFWSKSGWKITGCVLAS 801

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame.p2n.model -DRV=xlh
-Q=/abs/ABSSWEB.epool/US10600070/runac_21022006_112400_8653/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NOR=ext -HEADS=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abbs604
-USER=US10600070_@CGN_1_1_5315_@runac_21022006_112400_8653 -NCPU=6 -ICPU=3
-NO_MAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gsa1.*
10: gb_gsa2.*
11: gb_gsa3.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1761.5	43.4	2307	10	CL965374
2	1529	37.6	919	4	CNS09XTH
3	1038	25.5	741	4	CNS09XTH
4	1038	25.5	741	4	CNS09XTH
5	1002.5	24.7	813	7	CO071968
6	970	23.9	897	6	CD573714
7	881	21.7	785	9	B2437564

C	8	794.5	19.6	769	7	CO075595	CO075595 GR_Ba36G
C	9	789.5	19.4	871	10	DU050213	DU050213 14781.TG
C	10	788	19.4	832	7	CO079829	CO079829 GR_Ba42O
C	11	784	19.3	561	1	AI998415	AI998415 701545606
C	12	776	19.1	534	5	BQ834167	BQ834167 ALBT0128
C	13	774.5	19.1	703	7	CN918842	CN918842 030207ABP
C	14	767.5	18.9	741	8	CX543653	CX543653 UCRPT01_5
C	15	765	18.8	954	10	C9361431	C9361431 MBK3428TR
C	16	738	18.2	898	8	DR830109	DR830109 ZM_BPD007
C	17	733.5	18.1	631	5	BU046755	BU046755 PP_LBa002
C	18	720.5	17.7	849	7	CO117047	CO117047 GR_BD019
C	19	718.5	17.7	640	8	CX195205	CX195205 13-E01936
C	20	718	17.7	660	2	BI268376	BI268376 NF119C111
C	21	716.5	17.6	697	8	CX173275	CX173275 B01_69-11
C	22	694.5	17.1	728	8	CX674370	CX674370 UCRG510_6
C	23	671	16.5	611	8	CX117198	CX117198 B06_63-75
C	24	670	16.5	574	1	AW696905	AW696905 NF110C045
C	25	670	16.5	751	11	CR486014	CR486014 mch2-1550
C	26	652	16.0	820	8	DR972351	DR972351 ZM_BPD009
C	27	629	15.5	754	6	CD903230	CD903230 G356_109J
C	28	610.5	15.0	711	6	CR445760	CR445760 EST682105
C	29	609	15.0	646	3	BU464825	BU464825 B164825
C	30	605.5	14.9	593	7	CN904734	CN904734 010926ABD
C	31	604	14.9	922	7	CV264112	CV264112 WS02023.B
C	32	601.5	14.8	722	7	CO117046	CO117046 GR_BD019
C	33	596	14.7	664	8	DN939691	DN939691 6602_2.AE
C	34	590	14.5	819	7	CV241483	CV241483 MS02512.B
C	35	586	14.4	821	7	CK090561	CK090561 F013P64.3
C	36	573.5	14.1	723	7	CO079828	CO079828 GR_Ba42O
C	37	570	14.0	539	8	CX107243	CX107243 B113C23
C	38	563	13.9	866	6	CD573715	CD573715 UCRPT01_0
C	39	558	13.7	839	7	CK936162	CK936162 CGP100438
C	40	557.5	13.7	724	8	CK617940	CK617940 GABR1_36
C	41	550	13.5	746	8	CK617469	CK617469 UCRG510_6
C	42	548.5	13.5	521	7	CN914655	CN914655 030110ABN
C	43	548	13.5	768	8	CX543654	CX543654 UCRPT01_5
C	44	546	13.4	622	1	AV833644	AV833644 UCRPT01_5
C	45	544	13.4	415	3	BP785511	BP785511 BP785511

ALIGNMENTS

RESULT 1
LOCUS CL965374 2307 bp DNA linear GSS 21-SEP-2004
DEFINITION OIRFC012338 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL965374
VERSION CL965374.1 GI:52385433
KEYWORDS
SOURCE GSS.
ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 2307)
Ma,L., Wang,C., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

JOURNAL

COMMENT

Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1. 2307

ORIGIN

```

/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/notes="Oryza sativa exon trapped genomic sequences"

```

Alignment Scores:

Pred. No.:	2e-182	Length:	2307
Score:	1761.50	Matches:	390
Percent Similarity:	61.8%	Conservative:	119
Best Local Similarity:	47.4%	Mismatches:	213
Query Match:	43.4%	Indels:	101
DB:	10	Gaps:	17

US-10-600-070b-2 (1-801) x CL95374 (1-2307)

```

Oy 12 SerProPheGlnLeuCySarGLeuProProAlaThrThrLyLeuAArgSerH1Asn 31
    ::::::::::::::::::::
Db 40 GCGCCATTGCCCTTCCCTCCCTCCGCCGCCGCCGCCGCCGCCGCCGCCGCCCT 99
    ::::::::::::::::::::
Oy 32 ThrSerThrThrLeCyS--SerAlaSerLyTrpAlaAspArgLeuLeuSerAspHe 50
    ::::::::::::::::::::
Db 100 CACCCCTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCGCGCACTTC 159
    ::::::::::::::::::::
Oy 51 AsnPhe-----ThrSerAspSerSerSerSerPheAlaThrAlaThr 65
    ::::::::::::::::::::
Db 160 CACCTCTCCCTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 219
    ::::::::::::::::::::
Oy 66 ThrThrAlaThrLeuValSerLeuProProSerL1eAspArgProGluArgH1ValPro 85
    ::::::::::::::::::::
Db 220 GCGCCCTCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCGAGCGCGCGCAAGCTCCCTCCG 279
    ::::::::::::::::::::
Oy 86 IleProIleAspPheTyArgIn-----ValLeuGlyAlaGln 97
    ::::::::::::::::::::
Db 280 CTCCAAGTCGATTTCTTCAAGAGTGAGGCGCCCGTGAATCTGCTGCGGTCTTCAAGGCGCAG 339
    ::::::::::::::::::::
Oy 98 ThrHisPheLeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLyAspProPro 117
    ::::::::::::::::::::
Db 340 CCACATTTTCTTGGCGGATGCGATCAGAGGCGCTTCAGAGCAGCGATAGCCAGCCACCG 399
    ::::::::::::::::::::
Oy 118 GlnPheGlyPheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCyS 137
    ::::::::::::::::::::
Db 400 CAGTATGGCTACACGACGAGATGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
    ::::::::::::::::::::
Oy 138 GlnThrLeuSerAsnProArgSerArgArgGlnTyArgGlnGlnGlnGlnGlnGlnGlnGlnGln 157
    ::::::::::::::::::::
Db 460 GACACTCTCATGAACCGAAGCTCCGCGCATCAGATGATCGTGGCGCTTCTGAGAACCGT 519
    ::::::::::::::::::::
Oy 158 GlnAlaThrValIleThrAspValProTrpAspLyValProGlyAlaLeuCyValLeu 177
    ::::::::::::::::::::
Db 520 GAAGAAAGCTCTCACCATGATATTGCTTGGAGACAG----- 555
    ::::::::::::::::::::
Oy 178 GlnGlnGlyGlyGlnThrGlnIleValLeuArgValGlyGlnAlaLeuLeuLyGlnArg 197
    ::::::::::::::::::::
Db 556 ---GAGGCTGGGAGGACATCTGCTGTCTGTGAAGTGAAGACAGTGTCTTGGAGCTGG 612
    ::::::::::::::::::::
Oy 198 LeuProLySerPheLyGlnAspValValLeuValIleAlaLeuAlaPheLeuAspVal 217
    ::::::::::::::::::::
Db 613 CCACCCAAAGCGCTTCAAGCAGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
    ::::::::::::::::::::
Oy 218 SerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyArgInPheValGlu 237
    ::::::::::::::::::::
Db 673 TCAGGAGATCTTATGGCAGAAAGCTTCAGATGATTAATGGCTGCTCGAGAGGTCTCGAG 732
    ::::::::::::::::::::
Oy 238 GlnAlaLeuLySerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 257
    ::::::::::::::::::::
Db 733 AGGCTCTTCAAGCTCTTTCAGAGAGATGAGAGCAAGCATCTCGACCTGTATCTCTTCA 792
    ::::::::::::::::::::
Oy 258 GlnIleAspGlnThrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ::::::::::::::::::::
Db 793 CAGATTGATGAATCTCGAGAGATTAACACTCGCTGTGTATTTGAGAGCTTCTCTCCCTT 852
    ::::::::::::::::::::

```

```

Oy 278 ProLeuGlyAspAspTyAlaAlaLySarGLeuAenGlyLeuSerGlyValArgAsnIle 297
    ::::::::::::::::::::
Db 853 CCTATTGACACAGAGCATCATAGAAGCCGCAAGAGGCTTCAAGTGTGCGAGAACAT 912
    ::::::::::::::::::::
Oy 298 LeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlnLyS 317
    ::::::::::::::::::::
Db 913 TTGTGAGCGGTGGCAGAGAGATGTGTACCGTTGGAGAGAGATTTTCTGTGAAGCC 972
    ::::::::::::::::::::
Oy 318 PheMetAsnGlyAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAla 337
    ::::::::::::::::::::
Db 973 TTCATGACAGAGCTTTTGTAGATGATCATCATGTAACATGATGATTTCTTTCAAAA 1032
    ::::::::::::::::::::
Oy 338 ThrProSerAsnIleProAlaGlnSerPheGlnValTyGlnValAlaLeuAlaLeuVal 357
    ::::::::::::::::::::
Db 1033 ACACCGCAATGACATTCCTCTCGAATGTGTGAATTTACAAATGACACTTGACATCTGTC 1092
    ::::::::::::::::::::
Oy 358 AlaGlnAlaPheIleGlyLySerProH1sLeuLeuGlnAspAlaAspLySlnPheGln 377
    ::::::::::::::::::::
Db 1093 GCTCAAGCAATTATTAAGTAAAGCCCAATTATCATATGATGCGGATGATCTTTTGAA 1152
    ::::::::::::::::::::
Oy 378 GlnLeuGlnAlaLySerValMetAlaMetGlnIleProAlaMetLeuTyArgPThrArg 397
    ::::::::::::::::::::
Db 1153 CAGCTCAGAAAGTTCACATAGTCT-----CATTAAGCTTATGATATAT--- 1197
    ::::::::::::::::::::
Oy 398 AsnAsnTrpGlnIleAspPheGlyLeuGlnArgGlyLeuCyAlaLeuLeuIleGlyLyS 417
    ::::::::::::::::::::
Db 1198 -----GAGATGACCTTGTCATTGAAAGGCGCATTCGTGCTCATGCTAGTCGAGAGAT 1248
    ::::::::::::::::::::
Oy 418 ValAspGlnCyArgMetTrpLeuGlyLeuAspSerGlnAspSerGlnTyArgAsnPro 437
    ::::::::::::::::::::
Db 1249 GTTAGCAAGTGCAGAAATGTGCTGGAATGTAAATGATCTTACCATACAGAACGCC 1308
    ::::::::::::::::::::
Oy 438 AlaIleValGlnPheValLeuGlnAsnSerAsn---ArgAspAspAsnAspPheLeuPro 456
    ::::::::::::::::::::
Db 1309 AAAATTCTAGAGTTTATGTGACCAACTAGCATCAGAGAGAGATATGATCTTTCCA 1368
    ::::::::::::::::::::
Oy 457 GlyLeuCyAlaLeuLeuGlnGlnThrTrpLeuAlaGlyValValPheProArgPheArgAsp 476
    ::::::::::::::::::::
Db 1369 GGGCTGTGCAAGCTTTTGGAGACTTGCTGATCTTGAAGTTTCTTGAGACAGAGAT 1428
    ::::::::::::::::::::
Oy 477 ThrLyAspLySerPheLySlnLeuGlyAlaAspTyTyTrpAspAspProMetValLeuSerTy 496
    ::::::::::::::::::::
Db 1429 ACTCGGGGCGATGAGATTCAGATTCGATGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1488
    ::::::::::::::::::::
Oy 497 LeuGlnArgValGlnValValGlnGlySerProLeuAlaAlaIleAlaThrMetAlaArg 516
    ::::::::::::::::::::
Db 1489 CTAGAAAGGATGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1548
    ::::::::::::::::::::
Oy 517 IleGlyAlaGln-----HisValLyAsnAspAlaMetGlnAlaLeu 530
    ::::::::::::::::::::
Db 1549 CTGTGTGTCTCAAGCTACAGCTGCACTTGTAAGTGAATCAATGCTTATTCAGAGGTTTC 1608
    ::::::::::::::::::::
Oy 531 GlnLyValPhePro---SerArgTyTrpAspArgAsnSerIleGlnProLyAspVal 549
    ::::::::::::::::::::
Db 1609 AACAAAGTTTTCATTGATGAAACAGTTAGACAGGTGACGCCATGGAATAACT----- 1662
    ::::::::::::::::::::
Oy 550 GlnGlnThrValPheSerValAspProValGlyAsnAsnValGlyArgAspGlyGlnPro 569
    ::::::::::::::::::::
Db 1663 -----AAAGATGCG---CCT 1674
    ::::::::::::::::::::
Oy 570 GlyValPheIleAlaGlnAlaValArgProSerGlnAsnPheGlnThrAsnAspTyAla 589
    ::::::::::::::::::::
Db 1675 GGGGAGATCTT-----GAAAATTTTGACAG----- 1701
    ::::::::::::::::::::
Oy 590 IleArgAlaGlyValSerGlnSerSerValAspGlnThrThrValGlnMetSerValAla 609
    ::::::::::::::::::::
Db 1702 -----GAAAATGCACTGCT 1716
    ::::::::::::::::::::
Oy 610 AspMetLeuLySlnGlnAlaSerValLyIleLeuAlaIleGlyValAlaIleGlyLeuIle 629
    ::::::::::::::::::::
Db 1717 CATGATTCGAAATGCGCGCTTGAAATTAATCTCTGCTGCGCAGCTTTGCACTGTG 1776
    ::::::::::::::::::::
Oy 630 SerLeuPheSerGlnTyTrpPheLeuLySerSerSerSerPheGlnArgLyAspMet 649
    ::::::::::::::::::::

```

```

Db      1777 GCAGTAAATGCGGCCCAATAT-----TTGCCCTCGTAAGAGGCC 1815
      ::::|||||
Qy      650 ValSerSerMetGluSerAsp-----ValAlaThrIleGlySerValArgAlaAsp 666
      ::::|||||
Db      1816 CTTTCCTGCTATTGAGATGAGCATGATCTGTGGCACTTGCTAAATGATGCTGACTACT 1875
      ::::|||||
Qy      667 AspSerGluAlaLeu-----ProArgMetCaspAlaArgThrAla 679
      :|||
Db      1876 GATGATCTCTGCACCTAGTGAAGATCCAGTACATATCTCTAGAAATGATGCCAAGCTGGCA 1935
      :|||
Qy      680 GluAsnIleValSerIleTyrGlnIleIleYsSerLeuAlaPheGlyProAspPheArg 699
      :|||
Db      1936 GAAGATATTGTTCGCAAGTGGCAGAGTATCAATCTAAGGCTTGAGGACCGAAGACATTGC 1995
      :|||
Qy      700 IleGluMetLeuProGluValLeuAspGlyArgMetLeuYsIleTyrThrAspArgAla 719
      :|||
Db      1996 GTTCATCATGTCAGAGGTTCTTGATGGCAACATGCTTAAAGGTGGACTGACCGAGCA 2055
      :|||
Qy      720 AlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuYsLeuSerValAsp 739
      :|||
Db      2056 GCGGAGATTGAGCTCATGGGTGTTCTGGAGATACATCTATCCGATGTCAGCATTTGAT 2115
      :|||
Qy      740 SerValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSer 759
      :|||
Db      2116 AGCATCACTATCTCTCCAGATGCTGCAAGCGCATCTGTGAGGCTTACGATTGATGAGGCA 2175
      :|||
Qy      760 AlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThr 779
      :|||
Db      2176 GGGCAACTTACTGATGTTACTGACGCCAAGAACATGATTCATATGACACAAATTAACACT 2235
      :|||
Qy      780 ThrArgTyrGluValPheTyrSerIleSergIlyTyrPheIleThrGluGluSerVal 798
      :|||
Db      2236 ACCGGATGATGAGATGGCTTCTCCAGCTGAGAGGATGAGATACGAAAGGACAGTCT 2295
      :|||
Qy      799 IleuAlaSer 801
      :|||
Db      2296 CTCAGATCG 2304

RESULT 2
CNS09YJH
LOCUS
DEFINITION
CNS09YJH 919 bp mRNA linear HTC 04-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTJS892C08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX841670.1 GI:42406830
VERSION
BX841670.1
KEYWORDS
HTC; GSI; cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 919)
REFERENCE
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Queirer,F., Scarcelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 919)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

```

```

FEATURES
source
location/Qualifiers
1..919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTJS892C08"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..919
/gene="At5g42480"

ORIGIN
gene
US-10-600-070b-2 (1-801) x CNS09YJH (1-919)
Alignment Scores:
Pred. No.: 2,1e-157 Length: 919
Score: 1529.00 Matches: 301
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 37.6% Indels: 0
DB: 4 Gaps: 0

1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyAspArgPro 20
13 ATGGAAGCTTGATGATCGATCGATGATGCTCTCCCATTTCCATTAATGCCATTACCA 72
21 ProAlaThrThrIleValLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
73 CCGGCGACGACAAAGCTCCGACGATGACCAACACCTCTCAATCTGCTCCGACGAC 132
41 LysThrAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
133 AAATGGGCGCAGCGTCTCTCTCCGACTTAATTTCACTCCGATTCCTCTCTCTCC 192
61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
193 TTGGCCACCGCCACACACACGACCTCTGCTCTCCGACCATTAATGATGATGCC 252
81 GluArgHisValProIleProIleAspPheTyrGlnValIleGlyAlaGlnThrHisPhe 100
253 GAAGCGACGATCCCATCCCATCCCATTTGATTTACACAGGATTTAGAGCTCAACATTTTC 312
101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIysProProGlnPheGly 120
313 TTAAACCATGATATCAGAAAGATTCGAAGGTGAGGTTTGAACCGCGCAATTCGGT 372
121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
373 TTACGACGACACCGCTTTATACACCGGACGACATTTCTTAAGCTGCTTCCGAAACTCTG 432
141 SerAspProArgSerArgArgArgIlyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
433 TCTAATCTCGGCTGTAAGAGAGATCAATGAGAGGTCTTCTGATGATGAAGAGCTACA 492
161 ValIleThrAspValProTyrPheAspValProGlyValAlaLeuCyValLeuGlnGly 180
493 GTATCATCATATGTTCTTGAGTAAGGTTCTGAGTCTCTGATATTGCAAGAGGT 552
181 GlyIleThrGluIleValLeuArgValGlyValAlaLeuLeuYsGluArgLeuProIys 200
553 GGTGAGACTAGATAGTCTTCCGGTTGGTGGAGGCTCTGTTAAGAGAGGTTCCCTAAG 612
201 SerPheYsGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
613 TCGTTTAAGCAAGATGTTAGTTATGCGCTTGGCTTCTCGATGTCTCGAGGAGAT 672

```


QY 221 |AlaMetAlaLeuAbpProproApphe1eThcGlyTyrgluPheValGluGluAlaLeu 240
 |-----|
 Db 673 GCTATGGCATTGGATTCACCTGATTTATTAACGTGTTATAGTTTGTGGAAAGCTTGG 732
 QY 241 |LysLeuLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 260
 |-----|
 Db 733 AACCTTTTACAGGAGGAAAGAGCAAGTACCTTGGACCGGATTTACGTGCACAAATGAT 792
 QY 261 |GluThrLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 280
 |-----|
 Db 793 GAGACTTTGGAGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 852
 QY 281 |AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSer 300
 |-----|
 Db 853 GATGATTACCGCTGGAAAGACTAATGCTTTAAGCGGTGTGCGAATATTGTGTCT 912
 QY 301 |ValGly 302
 |-----|
 Db 913 GTTGA 918
 RESULT 3
 CNS09255
 LOCUS
 DEFINITION
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSTR1592A05 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 741)
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
 Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 741)
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction; Temple G.
 Genome members carried out sequencing and annotation; Castelli
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Piprap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?resource=Arabidopsis.
 FEATURES
 Source
 1..741
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /cdate="GSTR1592A05"
 /issue_type="Silique"
 /ecotype="Col-0"
 /plasmid="PCMVSPORT_6"
 1..741
 /gene="At5g42480"

ORIGIN

Alignment Scores:
 Pred. No.: 2,37e-103 Length: 741
 Score: 1038.00 Matches: 210
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 25.5% Indels: 0
 DB: 4 Gaps: 0
 US-10-600-070b-2 (1-801) x CNS09255 (1-741)
 QY 592 |AlaGlyValSerGluSerSerValaPgluThrThValGluMetSerValaAspMet 611
 |-----|
 Db 1 |GCTGGGCTCTCAGAGAGTAGCGTTGATGAACACTAGTGAATGTCCTGCTGATATG 60
 QY 612 |LeuLysGluLysAspValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeu 631
 |-----|
 Db 61 |TTAAGAGGCAAGTGTGAAGATCTAGCTGCTGGTGGCAATGAGATTCATCACTG 120
 QY 632 |PheSerGluLysTyrPheLeuLysSerSerSerPheGluArgLysAspMetValSer 651
 |-----|
 Db 121 |TTCAAGCAGAGATATTTCTTAAGACAGCTCATCTTTCAAGCAGATATGTTCT 180
 QY 652 |SerMetGluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeu 671
 |-----|
 Db 181 |TCTATGGAATCTGATGCGCTACATAGGGTCAGTCAGAGCTGACATTCAGAAAGCACTT 240
 QY 672 |ProArgMetAspAlaArgThrAlaGluAsnIleValSerLysThrGluLysIleLysSer 691
 |-----|
 Db 241 |CCAGAGATGATCTAGAGCTGAGAGAAATAGATCCAAAGGCGAGAAATTAAGTCT 300
 QY 692 |LeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMet 711
 |-----|
 Db 301 |CTGGCTTTGGCGCTGATCCACCCATAGAAATTTACAGAGGTTTGGATGGCGCAATG 360
 QY 712 |LeuLysIleThrThrAspArgAlaAlaGluThrAlaGluLeuValLysPyr 731
 |-----|
 Db 361 |CTGAAGATTTGAGACTGACAGAGAGCTGAAGATCCGACGCTGGTGTATGATAT 420
 QY 732 |ThrLeuLeuLysLeuSerValaPsserValThrValSerLysAspGlyThrArgAlaLeu 751
 |-----|
 Db 421 |ACACTGTGAACCTATCTGTGACAGTGCAGTGCACAGTCAAGTGAACCGGCTCTG 480
 QY 752 |ValGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGluAsnAsn 771
 |-----|
 Db 481 |GTGAAGCAACTCTGAGAGAGCTGCTGCTCTATCTGATTTGGTTCATCCAGAAACAAT 540
 QY 772 |AlaThrAspValArgThrTyrThrArgTyrGluValPheTyrSerLysSerGlyTyr 791
 |-----|
 Db 541 |GCTACGATGTCAAGACCTACACAAAGATAGGAATTTCTGCTCAAGTCAGGCTGG 600
 QY 792 |LysIleThrGluGlySerValaLysSer 801
 |-----|
 Db 601 |AAATCACTGAAGCTCTGTTCTTCATCA 630
 RESULT 4
 CNS09255
 LOCUS
 DEFINITION
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSTR151ZB09 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis
 1 (bases 1 to 741)
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
 Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,


```

Db      4 GTCTGCTCCGCGCAATGAATGGCCGAGCGCCCTCTGCTGATCTTCATTTCTCCCTGCC 63
Qy      56 SerserSerserSerserPhealathrAlathrThralathrleu---ValserleuPro 74
Db      64 CCCGATATATTCCTCTCTCTCTCTCTCTCTCTCAACCGCCACTCTTCTCCCTTACCTT 123
Qy      75 Proserle-----AsparProgluArghIsvAlProileProileasPheYrGln 92
Db      124 CCCCTTCTGCCCCCTTCTCTCCCGAAGCGCATGTTCCATTCCTCCCTGATTTCTCAAG 183
Qy      93 ValleuGlyValaglnthrhIsphleuThrasPglyYleArghalAphelGluAlaAr 112
Db      184 GTTTTAGAGCCGACATCTTTCTTAGTGATGATGATGATGATGATGATGATGATGATGATG 243
Qy      113 ValSerlyeProProglInphelIypheserAspAspAlaleuIleSerArghGlnIle 132
Db      244 GTTTCGAAACCGCCCTCAATATGAGTTCAATGACCAATATGACCAATATGACCAATAT 303
Qy      133 LeuGlnAlaAlaCyseGlnThrLeuSerAspProArghArghIyYrAsngIy 152
Db      304 CTCTAGCTGCTGCTGAAACCTATCTAACCTGCGCTTGAAGAAATTCACCAAGT 363
Qy      153 LeuLeuAspAspAglnGlnAlaThrValIleThrasPValProtrPAspIyValProgly 172
Db      364 CTGTGACGATGAGCGTGAACATCATCATCATCATCATCATCATCATCATCATCATCAT 423
Qy      173 AlaleuCyseValleuGlnIyGlyGlyThrGlnIleValleuArghValGlyGlnAla 192
Db      424 GCATGTGCGTGTGCAAGAGCTGGGAGACGTGAAGTGTGTTCAATTTGGGAGAGT 483
Qy      193 LeuLeuIyGlnIyArghLeuProIySerPheIyGlnIyAspValleuValMetAlaleu 212
Db      484 TTCTGTGAGAGAGAGCGTCCCAAGCATTTAAACAGATGTCGTTGGCAATGGCCCTG 543
Qy      213 AlAphleuAspValSerArghAspAlMetAlaleuAspPropAspPheIleThrGly 232
Db      544 GCTTATGTGACCTGTGAGAGGATGCTATGCTTGTGATCTCCAGATTTTCATTTGGTGT 603
Qy      233 TyGlnPheValGlnIyAlaleuIyLeuGlnIyGlnIyGlnIyAlaserSerleuAla 252
Db      604 TGTGAGTGTGCTTGAAGCGCTTGAACCTGTGCAAGAGAGAGTGCAGAGCTTGTCT 663
Qy      253 ProAspLeuArghAlaGlnIleAspGlnThrLeuGlnIyIleThrProArghYrValleu 272
Db      664 CCGGATTTGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
Qy      273 GluLeuLeuGlnIyLeuProLeuGlyAspAspTyAlAlaIyArghLeuAsngIyLeuSer 292
Db      724 GAACCTCTAGCCTTGGCCCTTGTGATGATGATGATGATGATGATGATGATGATGATG 783
Qy      293 GlyValArghAnIleuThrPserValGly 302
Db      784 GGTGTACGCAACATATTATGGCTGTGTGT 813

RESULT 6
CD573714      897 bp      mRNA      linear      EST 12-JUN-2003
LOCUS         UCRPT01_01_F12_T3 Poncirus trifoliata CTV-challenged cDNA library -
DEFINITION    UCR Poncirus trifoliata cDNA clone UCRPT01_01_F12, mRNA sequence.
ACCESSION     CD573714
VERSION       CD573714.1 GI:31669616
KEYWORDS      EST.
SOURCE        Poncirus trifoliata
ORGANISM      Poncirus trifoliata
REFERENCE     1 (bases 1 to 897)
AUTHORS       Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D.,
              Wananaker,S., Choi,Y. and Kinsan,T.
TITLE         Development of EST Resources and New Genetic Markers for California
              Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
JOURNAL       Unpublished (2003)

```

```

COMMENT       Contact: Mikelal Roose
              Department of Botany & Plant Sciences, University of California
              Riverside, CA, 92521-0124, USA
              Tel: 9097874137
              Fax: 9097874437
              Email: mikelal.roose@ucr.edu
              Seq primer: T3.

FEATURES
  source
    1..897
      Location/Qualifiers
        /organism="Poncirus trifoliata"
        /mol_type="mRNA"
        /cultiVar="PomeroY Op"
        /db_xref="taxon:37690"
        /clone="UCRPT01_01_F12"
        /issue_type="Phloem"
        /dev_stage="10 - 30 cm shoot"
        /lab_note="R. coli T0C121"
        /clone_lib="Poncirus trifoliata CTV-challenged cDNA
        library - UCR"
        /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
        Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
        greenhouse at University of California, Riverside. The
        action was a open-pollinated (very probably selfed)
        seedling of Poncirus trifoliata cv PomeroY that was
        rootstock was sweet orange infected with citrus tristeza
        virus (CTV) isolate T514 over 1 year before sampling (CTV
        infects sweet orange, but not genotypes carrying the CTV
        resistance gene. Shoots 10-30 cm long were harvested in
        October 2000, and the green phloem (bark) was removed and
        frozen quickly in dry ice. Total RNA was extracted using
        Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
        library was made, and 0.5 million primary lambda cDNA
        clones were in vivo excised to give a population of
        pluescript SK(-) phagemids. All steps to this point were
        performed in the M. Roose lab at the University of
        California, Riverside by X. Ye. Phagemids were plated,
        plasmid DNA purified, cDNA clones archived, and DNA
        sequences determined bi-directionally using an ABI3730 at
        the University of California Riverside Genomics Institute,
        Core Instrumentation Facility, (Choi, Kinsan).
        Chromatogram files were transmitted to UC Riverside (by
        Choi), then processed at UC Riverside (by Wananaker) using
        the HarVest pipeline (http://harvest.ucr.edu) to remove
        vector and cloning oligo sequences and various
        contaminants, and to trim to a high quality region.
        Sequences that retained a phed 17 region of at least 100
        bases were deposited to GenBank."

ORIGIN
Alignment Scores:
Pred. No.:      1,04e-95      Length:      897
Score:          970.00      Matches:      189
Percent Similarity: 81.2%      Conservative: 32
Best Local Similarity: 69.5%      Mismatches: 51
Query Match:    23.9%      Indels:      0
DB:             6      Gaps:      0

US-10-600-070b-2 (1-801) x CD573714 (1-897)

Qy      248 AlaSerSerleuAlaProAspLeuArghAlaGlnIleAspGlnThrLeuGlnIyIleThr 267
Db      80 GCCAGTAGTCTCGCTCCAGATTGTCAGGACAGATGACGATGCTGGAGGAGATCAAC 139
Qy      268 ProArghYrValleuGlnIyLeuProLeuGlyAspAspTyAlAlaIyArghLeuAsngIy 287
Db      140 CCACGCTGTGTTCTGGAACCTTTAGGCTTAAGTGTAGTGTCAAGCAAGAGA 199
Qy      288 LeuAsngIyLeuSerGlyValArghAnIleuThrPserValGlyGlyGlyAlaSer 307
Db      200 GAAGAAAGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 259
Qy      308 AlaleuValGlyGlyLeuThrArghIyGlyPheMetAsngIyAlaPheLeuArghMetThr 327

```

```

Db      260 GCATTTGCTGGGGGATTACAGCGAATCTTCAATGATGAGGCTTTCTACGATGACA 319
Qy      328 AAlaIaGluGlnValaAspLeuPheValAlaThrProSerAnIleProAlaGluSerPhe 347
Db      320 TCACCTGAGCGAGTTAACTATTTTCTGCGCACCAACCAATACATCCCGCTGAAACTTTT 379
Qy      348 GluValTyrGluValAlaLeuValaLeuValaIleGlnIlePheIleGlyValaProHis 367
Db      380 GAGCTTATGAGAGTGGCACTTGCTTGCTCCCAAGCTTTGGGGTAGCAACCTCAT 439
Qy      368 LeuLeuGlnAspAlaAspLysGlnPheGlnIleLeuGlnIleAlaLysValaMetAlaMet 387
Db      440 CTCATTGAGATGCTGATTAACATGTTCAAGCACTTCAAGCAATTAAGTACCAAGCTTTA 499
Qy      388 GluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGlu 407
Db      500 CGGACCTTGCTGCTATCTATATTCCTTGGAATAACATGAGATGAGTTCGCAATGGAA 559
Qy      408 ArgGlyLeuCyValaLeuLeuIleGlyValaValaAspGluCyArgMetTrpLeuGlyLeu 427
Db      560 AGGGCTATGATTCACGCTTGATGAGTAAAGTATGATGATGCTTATGATGATGATGATGAT 619
Qy      428 AspSerGluAspSerGlnTyrArgAsnProAlaIleValaGluPheValaLeuGluAsnSer 447
Db      620 GATAGTATAGTACCTCCCATACAGAAATCCACCATTTGATTTGTTTGGAAACTCA 679
Qy      448 AsnArgAspAspAsnAspAspLeuProGlyLeuCyValaLeuGluTrpLeuAla 467
Db      680 AAGCAACTATATGACATGATCTTCTGCTCTGTAAGCTGTAGAGACATGCTGCA 739
Qy      468 GlyValaIlePheProAlaGlyPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr 487
Db      740 GAGGTGTTTCCCTAGATTCAGGACACAGATATATACGTTCAAACTTGATGATAC 799
Qy      488 TyrAspAspProMetValaLeuSerTyrLeuGluValaGluValaIleGlnIleSerPro 507
Db      800 TATGATATCTACTGCTCTTCCGATATTTGAAAGATCGAGGGAATGATGATTCGCC 859
Qy      508 LeuAlaAlaAlaAlaThrMetAlaArgTleGlyAla 519
Db      860 TTAGCTGACAGACCTGCATAGTAGGCTTGAGCT 895

RESULT 7
BZ437564/c 785 bp DNA linear GSS 13-DEC-2002
LOCUS BONRN72TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONRN72,
DEFINITION genomic survey sequence.
ACCESSION BZ437564.1 GI:26691135
VERSION BZ437564.1 GI:26691135
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
TITLE 1 (bases 1 to 785)
JOURNAL Azele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
PUBMED Uteback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
COMMENT Whole genome shotgun sequencing of Brassica oleracea and its
GENE application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Other_GSSs: BONRN72TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

```

```

FEATURES
    source
        location/Qualifiers
            1..785
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000D3"
            /db_xref="taxon:3712"
            /clone="BONRN72"
            /clone_1kb="BO.1.6.2 KB tot"
            /note="vector: pHD51; Site 1: BstXI, 1.6-2 kb sheared
            total DNA inserted into pHD51 using BstXI linkers"

ORIGIN
Alignment Scores:
Pred. No.: 5,59e-86 Length: 785
Score: 881.00 Matches: 191
Percent Similarity: 83.5% Conservative: 6
Best Local Similarity: 80.9% Mismatches: 4
Query Match: 21.7% Indels: 35
DB: 9 Gaps: 2

US-10-600-070B-2 (1-801) x BZ437564 (1-785)

Qy      169 LysValProGlyAlaLeuCyValaLeuGlnGluGlyValGluTrpGluIleValaLeuArg 188
Db      693 AAGGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
Qy      189 ValGlyValaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 208
Db      633 GTAGAGAACCTTGTCTTAAAGAGAGGCTTCAAGCTTCAAGCAAGATGTGTTTGG 574
Qy      209 ValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProAsp 228
Db      573 GTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514
Qy      229 PheIleThrGlyTyrGluPheValaGluGluAlaLeuLysLeuLeuGlnGlnGluGlyAla 248
Db      513 TTTATTACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
Qy      249 SerSerLeuAlaProAspLeuArgAlaGlnIleAspGluTrpLeuGlnGluIleThrPro 268
Db      453 AGCAGCTTCCACCTGATTTAAGCGCCCAATGATGATGATGATGATGATGATGATGATGATGAT 394
Qy      269 Arg-TyrValLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 288
Db      393 CGGTTATGATTTGAGCTCTTCTGTTTAACTCTGCGGAGATGAT-----AAACACA 343
Qy      288 uAsnGlyLeuSerGlyValaArgAsnIleLeuTrpSerValGlyGlyGlyAlaSerAl 308
Db      342 AGATGCTTTAAGCGGTGTACGAATATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283
Qy      308 AlaValaGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLeuArgMetThrAl 328
Db      282 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 223
Qy      328 AlaGluGln----- 331
Db      222 ATCTGAGCAAGCTGTAACCCCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 163
Qy      332 -----ValaAspLeuPheValaAlaThrP 339
Db      162 TGGTTTCCCATCTTATCTCTGCTCTTGTGTCATAGTTGATCTTTTGTGCTACCC 103
Qy      339 roSerAnIleProAlaGluSerPheGluValaTyrGluValaAlaLeuAlaLeuValaIleG 359
Db      102 CGAGTATATTTCCAGCAGAGTCAATTGAAATTATGAAATTGCGCTGGCGCTTGCTGCTC 43
Qy      359 IlnAlaPheIleGlyLys-LysProHisLeuLeuGlnAspAla 372
Db      42 AAGCTTTATAGTAAATTAACCCCACTCTTACAGAGCC 1

RESULT 8
LOCUS CO075595 769 bp mRNA linear EST 15-JUN-2004

```

DEFINITION GR_Ea36G17.r GR_Ea Gossypium raimondii cDNA clone GR_Ea36G17.3', mRNA sequence.

ACCESSION CO075595

VERSION CO075595.1 GI:48745076

KEYWORDS EST

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 769)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C., and Wang, R. A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 36 row: G column: 17.
Location/Qualifiers

FEATURES

source 1..769
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea36G17"
/cruise_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH108"
/clone_lib="GR_Ea"
/note="Vector: pCMV.SPORT-6.1. Site 1: Not; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into Not-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.:	1,92e-76	Length:	769
Score:	794.50	Matches:	157
Percent Similarity:	79.9%	Conservative:	34
Best Local Similarity:	65.7%	Mismatches:	41
Query Match:	19.6%	Indels:	7
DB:	7	Gaps:	4

US-10-600-070b-2 (1-801) x CO075595 (1-769)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProphGlnLeuCyArgLeu--- 19

DB 54 ATGAAATCTTGGAGGACATTAGCATGGGCTGGACCTCTTACACTTACTCTTTATTC 113

QY 20 ---ProProAlaThrThrIleuArgArgSerHisAsnThrSerThrThrIleCysSer 38

DB 114 CACCTCTCTAAACCTCGACATCCAC-----CACGTTCAACTACCGTGTGCTCC 167

QY 39 AlSerSerThrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSer 58

DB 168 GCCAGTAAATGGCGCGAGCGCTCTTGTGACTTCCCAATTCTCCCTGCCCGCATAT 227

QY 59 SerSerPheAlaThrAlaThrThrAlaThrIleu---ValSerLeuProProSerIle 77

DB 228 TCGCTCTCTCTTCTCTCTCAACGCACTTTCTCCCTTACCCCTTACCCCTTCTC 287

QY 78 -----AspArgProGluArgHisValProIleProIleAspPheArgValValLeuGly 95

DB 288 GCCCTTCTCTCTCCGCAACGCGCATTTCCATTCCTTGAATTTCTCAAGGTTTGA 347

QY 96 AlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIys 115

DB 348 GCCGAGACTATTCTTAGTGATGATGATCAGAGAGCGCTATGAAGCAAGGTTTCGAA 407

QY 116 ProProGlnPheGlyPheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnAla 135

DB 408 CCCTCTCAATATATGGTTCAAGTCAAGACACCATATATAGCGAAGACAGATTCTTAGCT 467

QY 136 AlaCysGlnThrIleuSerAsnProAspSerArgArgGlnIleArgGlnIleLeuAsp 155

DB 468 GCGTGTGAAGAACCTTACTTAACTCGCTGTAGCGAAATTAACAACAGGCTTGTGAC 527

QY 156 AspGluGluAlaThrValIleThrAspValProThrAspIleValProGluAlaLeuCy 175

DB 528 GATGACCGTACACTTATCATCTCAAGTCCCTGGGACAGAGTTCTTGAGCATGTGC 587

QY 176 ValLeuGlnGluGlyIleGlnIleValIleuArgValGlyGluAlaLeuLeuIys 195

DB 588 GTGTTCAAGAACTGCGGGAGACTGAAGTGTGCTTCAATTTGGGAGATTGTGAGA 647

QY 196 GluArgLeuProIleSerPheIysGlnAspValValLeuValMetAlaLeuAlaPheLeu 215

DB 648 GAGAGGCTGCGCCCAAGCATTTAAACAGATGTGTTCTGGCAATGGCCCTGCTTATGTG 707

QY 216 AspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyIle 234

DB 708 GACTGTGCAAGGATGCTATGCTTGTGATCTTCCAGATTTCATTGTGTGTGTGAG 764

RESULT 9

DU050213/c 871 bp DNA linear GSS 12-AUG-2005

LOCUS DU050213

DEFINITION 144781 Tomato HindIII BAC library Lycopersicon esculentum genomic clone LE_HBA0013B21 5, genomic survey sequence.

ACCESSION DU050213

VERSION DU050213.1 GI:72480609

KEYWORDS GSS.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 871)
Mueller, L.A., Buehl, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J., Van Eck, J. and Stack, S.
BAC end sequencing from three Solanum lycopersicon libraries
Unpublished (2005)
Other GSSs: 144780
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Insert length: 76680 Std Error: 0.00
Plate: 13 row: B column: 21
Seq primer: T7
Class: BAC ends
High quality sequence start: 43
High quality sequence stop: 789.
Location/Qualifiers

FEATURES

source 1..871
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultiivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="LE_HBA0013B21"
/lab_host="E. coli"
/clone_lib="Tomato HindIII BAC library"
/note="Vector: pBeloBAC11; Site_1: HindIII"

ORIGIN

Alignment Scores:

Pred. No.:	8.35e-76	Length:	871
Score:	789.50	Matches:	171
Percent Similarity:	70.1%	Conservative:	17

Best Local Similarity: 63.8% Mismatches: 40
 Query Match: 19.4% Indels: 40
 DB: 10 Gaps: 4
 US-10-600-070b-2 (1-801) x DU050213 (1-871)

```

Qy 234 GluPheVal-----GluGluAlaLeuYsLeu---LeuGln 244
    |||||
Db 843 GAATTTCATAGCTTACAGTGAAGCAATATATAGAGAACTTCTTCCCTTTTCCAG 784

Qy 245 GluGluGluValAspSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGlu 264
    |||||
Db 783 GAAGAAAGTCCAAATATCTTGCCCTTGATCTGCATCCAGATCCAGATAGATGAGACATTGAA 724

Qy 265 GluIleThrProArgTyrValIleLeuGluLeuGluProLeuGluYAspAspTyrAla 284
    |||||
Db 723 GAATTAATCCAGCTACGTAAGTGAACCTTCTTCTTGATGAAATACCGA 664

Qy 285 AlaIleArgLeuAsnGluLeuSerGlyValArgAsnIleLeuTyrSerValGlyGly 304
    |||||
Db 663 ATGAAAAGAGTACAGGCTCTTCAAGTGTGCGCAACTTTGTGGCTGTGAGAGAGA 604

Qy 305 GlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLeu 324
    |||||
Db 603 GGAGCAGCTCAATTTTCAGGGGGGTTCAACAGAGAAATTTTCAATGATGAGGCTTCTTA 544

Qy 325 ArgMetThrAlaAlaGluGln----- 331
    |||||
Db 543 CGGATGACAGCTCTGAGCAGGTGGCTTAATTCAATTATTTTGGTTCTTTTGT 484

Qy 332 -----ValAspLeuPh 335
    |||||
Db 483 AACTGTTCTTTCACCTCTCTCTTCAATTTCTCTTCTTTTAAAGGTGACCTCTT 424

Qy 335 eValAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAl 355
    |||||
Db 423 CGTCGCAAGCCCAAGTAACTTCTGCAAGAACTTTGAAAGTTTACGGGGTGGGCTTGC 364

Qy 355 AlaLeuValAlaGlnAlaPheIleGlyLysProHisLeuLeuGlnAspAlaAspLysGlu 375
    |||||
Db 363 ACTGTGTGCTCAAGCTTTTGGGAAAAAACCTCATCTCATTCAMATCTGATTAACCT 304

Qy 375 nPheGlnGlnLeuGlnAlaIleValysValIleAlaMetGluIleProAlaMetLeuTyrAs 395
    |||||
Db 303 TTTTCAGCAGCTTTCAGCAGACCAAGATTACAGCTTACGCACTTCTGTCTGATACAC 244

Qy 395 pThrArgAsnAsnTyrpGluIleAspPheGlyLeuGluArgGlyLeuCyAsAlaLeuIle 415
    |||||
Db 243 TGTTAGAGAAACCGTGAATATAGCTTCGCTTGGAGAGGGGCTTTGTTCACTGCTGT 184

Qy 415 eGlyLysValAspGluCyAsArgMetTyrPheGlyLeuAspSerGluAspSerGlnTyrAr 435
    |||||
Db 183 TGAAGAAAGTCAAGTCAAGTGTGCTTGGTGGGCTTGAGACAGTGAAGACCTCCCTTAAG 124

Qy 435 GAsnProAlaIleValIleGluPheValIleGluAsnSerAsnArgAspAsp---AsnAspAs 454
    |||||
Db 123 AGATCCAGCTATAGTGACTTTTGTTCAGAAACCTCAAGAGACGACAAAGAAATGATCT 64

Qy 454 pLeuProGlyLeuCyAsLysLeu 461
    |||||
Db 63 GCTCCCTGACCTATGTAAGCTT 42

RESULT 10
LOCUS CO079829 832 bp mRNA linear EST 15-JUN-2004
DEFINITION GR_Ea42007.r GR_Ea Goseypium raimondii cDNA clone GR_Ea42007.3',
mRNA sequence.
ACCESSION CO079829
VERSION CO079829.1 GI:48749310
KEYWORDS EST.
SOURCE Goseypium raimondii
ORGANISM Goseypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Goseypium.
 1 (bases 1 to 832)
 Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.
 Global assembly of Cotton ESTs
 Unpublished (2004)
 CONTACT: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 Plate: 42 row: 0 column: 07.
 Location/Qualifiers
 1..832
 /organism="Goseypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Ea42007"
 /isue_type="whole seedlings"
 /dev_stage="first true leaves"
 /lab_host="DH10B"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

Alignment Scores:
 Pred. No.: 1.14e-75 Length: 832
 Score: 788.00 Matches: 172
 Percent Similarity: 67.6% Conservative: 30
 Best Local Similarity: 57.5% Mismatches: 61
 Query Match: 19.4% Indels: 36
 DB: 7 Gaps: 5

```

US-10-600-070b-2 (1-801) x CO079829 (1-832)

Qy 343 ProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIle 362
    |||||
Db 3 CTTGCAAGAAAGTTTCAAGTTTATGCTGTAAGCACTTGTCTCAAGCTTTTATTA 62

Qy 363 GlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnAla 382
    |||||
Db 63 AACAGAAACCTCATCTCAATTAAGATGCTGATTAACCTTCCAGCAGCTTACAGACT 122

Qy 383 LysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTyrpGluIle 402
    |||||
Db 123 AAGTAAACAACCTTTAGAAACAGTGTCTCTTATGCTCTGTCGGAACCGTGAATA 182

Qy 403 AspPheGlyLeuGluArgGlyLeuCyAsAlaLeuIleGlyLysValAspGluCyAsArg 422
    |||||
Db 183 GACTTGTCTTGAAGAGGGCTCTCTTCACTTCTGCTGGGAGACTTATAGTGCTGCT 242

Qy 423 MetTyrPheGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValIleGluPhe 442
    |||||
Db 243 TCGTGGTGGCTTACAGTGAAGTACTCCCTTATAGAAATCATCTATTTGTAATTT 302

Qy 443 ValLeuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyAsLysLeuLeu 462
    |||||
Db 303 GTCTGGAACAACTCAAGATGACATGACAGATCTTCCCGGCTTTGCAAACTGCTG 362

Qy 463 GluThrTyrPheLysAlaGlyValIlePheProArgPheArgAspThrLysAspLysPhe 482
    |||||
Db 363 GAGGCATGGCTATAGAGATGATTTTCTTCAATTAAGACCAAAAGATATACATTC 422

Qy 483 LysLeuGlyAspTyrTyrAspAspPrometValLeuSerTyrLeuGluArgValGluVal 502
    |||||
Db 423 AACCTTGAAGATTATATGATGATCTCACTGCTCCTGAGATATTTTAAAGGCTTGAGGGA 482

```


Fax: 919 515 1695
Email: mbarrie@unc.edu
Plate: 1 row: H column: 5
Seq primer: T3.

FEATURES

Source

Location/Qualifiers
1..534
/organism="Arabidopsis lyrata"
/mol_type="mRNA"
/cultivar="Karumaki"
/db_xref="taxon:59689"
/clone="PIWI-D03"
/tissue_type="Inflorescence"
/clone_lib="Arabidopsis lyrata Inflorescence pCMV-PCR library"
/note="Vector: pCMV-PCR (Stratagene); Created using PCR Library Construction Kit (Stratagene)"

ORIGIN

Alignment Scores:

Pred. No.:	1,17e-74	Length:	534
Score:	776.00	Matches:	155
Percent Similarity:	95.7%	Conservative:	2
Best Local Similarity:	94.5%	Mismatches:	7
Query Match:	19.1%	Indels:	0
DB:	5	Gaps:	0

US-10-600-070b-2 (1-801) x BQ834167 (1-534)

```
QY 638 LeuLYSerSerSerSerPheGlnArgValAspMetValSerMetGluSerAspVal 657
    |||||:::|
DB 13 CTCGAAAGGACGCTCATCTTTCAACGACGATATGGCTTCTTATGAAATCGACATT 72
QY 658 AlaThrIleGluSerValArgAlaAspAspSerGlnAlaLeuProArgMetAlaArg 677
    |||||:::|
DB 73 GCTACCATAGGGTCAGTGAAGACCTGATGATTCAGAACCTTCCAGAAATGATGCTAGG 132
QY 678 ThrAlaGluAsnIleValSerLySTpGlnLysIleLysSerLeuAlaPheGluProAsp 697
    |||||:::|
DB 133 ACTGCGACGAGTATGATTCMAATGSCAGAAATCAAGTCTCAGGCTTTGGGCAATGAT 192
QY 698 HisArgIleGluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTPThAsp 717
    |||||:::|
DB 193 CACTGATATGAATATGCTACAGAGGTTTGGATGGGGAATGCTGAAGATTGGAGTAC 252
QY 718 ArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuLysSer 737
    |||||:::|
DB 253 AGAGCTGTGGAAACCTGGCGAGCTTGGCTGTATATATACACTGTGAACATATCT 312
QY 738 ValAspSerValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGlu 757
    |||||:::|
DB 313 GTTGACAGTGTGACGATCTCAGACGATGAAACCGGCTGCTGGTGAAGCACTTGGAG 372
QY 758 GluSerAlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThr 777
    |||||:::|
DB 373 GAGTCTGCTTGTATCTGATTTGGTTTCATCCAGAAACATGCTACAGATGTCAGAAC 432
QY 778 TyrThrThrArgTyrGluValPheTTPSerLysSerGlyTTPlysIleThrGluLysSer 797
    |||||:::|
DB 433 TACACACACAGATAGCAAGTGTCTGCTCAAGTCAAGGTGGAATCACTGAAGGCTCT 492
QY 798 ValLeuAlaSer 801
    |||||:::|
DB 493 GTTCTGGCATCA 504
```

RESULT 13
CN918842
LOCUS 030207ABPB006290HT (ABPB) M9 root tips Malus x domestica cDNA clone
DEFINITION ABPB006290, mRNA sequence.
ACCESSION CN918842
VERSION CN918842.1 GI:48391655
KEYWORDS EST.
SOURCE Malus x domestica

ORGANISM

Malus x domestica

Malvaceae: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE

1 (bases 1 to 703)

AUTHORS

Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,

TITLE

McArdney, S., Newcomb, R., Rose, G., Snowden, K., Walton, E. and Yauk, Y.

JOURNAL

HortResearch Apple EST Project

COMMENT

Unpublished (2004)

CONTACT

Contact: Gleave, A.

Sequencing Facility

The Horticulture and Food Research Institute of New Zealand Ltd

Address

120 Mt Albert Rd, Mt Albert, Auckland, New Zealand

Tel:

Tel: 00 64 09 815 4200

Fax:

Fax: 00 64 09 815 4201

Email:

Email: est@hortresearch.co.nz.

Library

Library sequenced by
Genesis Research & Development"

ORIGIN

Alignment Scores:

Pred. No.:	2.67e-74	Length:	703
Score:	774.50 <td>Matches:</td> <td>156</td>	Matches:	156
Percent Similarity:	74.9% <td>Conservative:</td> <td>26</td>	Conservative:	26
Best Local Similarity:	64.2% <td>Mismatches:</td> <td>34</td>	Mismatches:	34
Query Match:	19.1% <td>Indels:</td> <td>27</td>	Indels:	27
DB:	7	Gaps:	4

US-10-600-070b-2 (1-801) x CN918842 (1-703)

```
QY 1 MetGluAlaLeuSerHisValGlyIleGly-----LeuSerProPheGln 15
    |||||
DB 14 ATGAAACCTTCMAACATTCGGCATCGGCTTCTACATCCCGGCTTCCTTCCTTCGCG 73
QY 16 LeuCysArg-----LeuProPheAlaThrThrLysLeuArgSerHisAsn 31
    |||||
DB 74 CACCACCGAAACCCAAAGCTCCTCC-----SerPheAlaThrAlaThrThr 103
QY 32 ThrSerThrThrIleCysSerAlaSerLySTpAlaAspArgLeuSerAspPheAsn 51
    |||||
DB 104 -----ACCATCTGCTTCGCGCAGCAATGCGGAGGCTCCTCGCGCACTTCCAA 154
QY 52 PheThrSerAspSerSerSer-----SerPheAlaThrAlaThrThr 66
    |||||
DB 155 TTCCTCGGAGACCTTCCTCCTCCTCCTCAGACCAACATTCCTCCTCGGACCGCC 214
QY 67 ThrAlaThrLeuValSerLeuProPheSerIleAspArgProGluArgHisValProIle 86
    |||||
DB 215 ACTCTGCTCCTCCCACTGCTCCTCCGCAATTTCTCCCGGAGGCCAGCTGTCCGTC 274
QY 87 ProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPheLeuThrAspGlyIleArg 106
    |||||
DB 275 CCCATCGATTTTACACAGGTGCTGGCGCCAGCAGCAATTTCTCGGAGCGGATATAGG 334
QY 107 ArgAlaPheGluAlaArgValSerLySTpProGluPheGlyPheSerAspAlaLeu 126
    |||||
DB 335 AGGCGCTAGAGGCTGAGGCTTCAGACCGCTCAGTACGCTTACCCAGAGGCGCTG 394
QY 127 IleSerArgArgGlnIleLeuGlnAlaIleValCysGluThrLeuSerAspProArgSerArg 146
    |||||
DB 395 TTACGCCGAGGAGGATCTTCTAGCGGCTTGGGAAACCTTACCGACCCCGCTCAGA 454
QY 147 ArgGluTyrAsnGluLysLeuLeuAspArgGluAlaThrValIleThrAspValPro 166
    |||||
DB 455 AGAGAGTACACACAGAGCCTTGGCGAAGAGGAGATGAAACCATCATCACTCAAGTTCTT 514
```

QY 167 TTPAEPlyValProglValaleuCyValleuGluGlyGlyGluhrGluVal 186
 DB 515 TGGGATPAAGCTTCTGAGCTGTGCTGCTGCAAGAGCTGGAGACTGAGCTGGT 574
 QY 187 LeuArgValIGlyGlualeuLeuLySGlyArgLeuProLySerPheLySGlnAepVal 206
 DB 575 CTTCAAAATTTGGGAGAGTTTGGCTAAGAGAGGCTCCCAAGTCGTTCAAGCAAGATGTC 634
 QY 207 ValLeuValMeAlaleuAlaPheLeuAepValSerArgAapAlaMeAlaleuAapPro 226
 DB 635 GTTTCGTCATGCGACCTGCTTATGTTGACATGTGAGGATGCAATGCAATGTCCTCCG 694
 QY 227 ProAepPhe 229
 DB 695 CCTGATTTT 703
 RESULT 14
 CX543653 741 bp mRNA linear EST 12-JAN-2005
 LOCUS UCRPT01_5_008.C05.T3 Poncirus trifoliata CTV-challenged cDNA
 DEFINITION library - UCRPT01-UCR2 Poncirus trifoliata cDNA clone
 UCRPT01_008.T3_C05, mRNA sequence.
 CX543653
 ACCESSION CX543653.1 GI:57570678
 VERSION EST.
 KEYWORDS Poncirus trifoliata
 SOURCE Poncirus trifoliata
 ORGANISM Poncirus trifoliata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eustosids II; Sapindales; Rutaceae; Poncirus.
 1 (bases 1 to 741)
 ROOSE, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
 Wamamaker, S., Choi, Y. and Kington, T.
 Development of EST Resources and New Genetic Markers for California
 Citrus - Poncirus trifoliata CTV-challenged phloem - UCRPT01-UCR2
 Unpublished (2004)
 CONTACT: Mikeal Roose
 Department of Botany & Plant Sciences, University of California
 Riverside, CA, 92521-0124, USA
 Tel: 9097874137
 Fax: 9097874437
 Email: mikeal.roose@ucr.edu
 Seq primer: T3
 Location/Qualifiers
 1..741
 /organism="Poncirus trifoliata"
 /mol_type="mRNA"
 /cultivar="Pomeroiy OP"
 /db_xref="taxon:37690"
 /clone="UCRPT01_008.T3.C05"
 /tissue_type="bark (with phloem)"
 /dev_stage="10 - 30 cm shoots"
 /lab_host="E. coli TJC121"
 /clone_lib="Poncirus trifoliata CTV-challenged cDNA
 library - UCRPT01-UCR2"
 /note="Vector: lambda Uni-ZAP XR, excised phagemid;
 Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
 greenhouse at University of California, Riverside. The
 section was an open-pollinated (very probably selfed)
 seedling of Poncirus trifoliata cv Pomeroiy that was
 selected as homozygous for the CTV resistance gene. The
 rootstock was sweet orange infected with citrus tristeza
 virus (CTV) isolate TS14 over 1 year before sampling (CTV
 infects sweet orange, but not genotypes carrying the CTV
 resistance gene. Shoots 10-30 cm long were harvested in
 October 2000, and the green phloem (bark) was removed and
 frozen quickly in dry ice. Total RNA was extracted using
 RNeasy reagent (Qiagen). Poly(A) RNA was purified, a cDNA
 library was made, and 0.5 million primary lambda cDNA
 clones were in vivo excised to give a population of
 phagescript SK(-) phagemide. All steps to this point were
 performed in the ML Roose lab at the University of

California, Riverside by X. Ye. Phagemids were plated,
 plasmid DNA purified, cDNA clones archived, and DNA
 sequences determined bi-directionally using an ABI3730 at
 the University of California Riverside Institute of
 Integrative Genome Biology Genomics Core Instrumentation
 Facility, (Choi, Kington). Chromatogram files were
 downloaded by FTP by Close, then processed by Wamamaker
 (Close lab) using the Harvest pipeline
 (http://harvest.ucr.edu) to remove vector and cloning
 oligo sequences and various contaminants, and to trim to a
 high quality region. Sequences that remained a phred 17
 region of at least 100 bases were assembled, then chimeras
 were removed following manual inspection of assemblies
 (Close, Roose, Wamamaker). Sequences that survived all
 removal steps were submitted to GenBank."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.73e-73 Length: 741
 Score: 767.50 Matches: 152
 Percent Similarity: 77.2% Conservative: 31
 Best Local Similarity: 64.1% Mismatches: 49
 Query Match: 18.9% Indels: 5
 DB: 8 Gaps: 2
 US-10-600-070B-2 (1-801) x CX543653 (1-741)
 QY 1 MetGluAlaLeuSerHisValGlyLeuSerProPheGlnLeuCyAargLeuPro 20
 DB 32 ATGCAAGCTTGGAGACATGATGTTGGTCTTCAAGTCCAGCTGCTCCAGGCGCT 91
 QY 21 ProAlaThrThyLeuValArgArg-----SerHisAsnThrSerThrTyrLeuCy 37
 DB 92 CAGCCTCCATAGAGATGAAGAACCTCTATTTCCAAATGAAATGACACCACTGCTCC 151
 QY 38 SerAlaSerLygTrpAlaAparGluLeuSerAapPheAsnPheserAapSerSer 57
 DB 152 TCGGCGAGCAAAATGGGCGCCACCGCTCCGCGACCTTCATTCACCAACCGCGACAT 211
 QY 58 SerSerSerPheAlaThrAlaThrThrAlaThrLeuValSerLeuProPheSerile 77
 DB 212 TCCTCTCTCTCTCTATCTTCAACACCAACCGTACCTTCAACCTCCCTCTCTACT--- 268
 QY 78 AspArgProGluArgHisValProIleProIleAapPheTyrGlnValleuGlyAlaGln 97
 DB 269 ---CCACCAACCGCGACGATCCATCCCATGATCTTATGAGCGGTGGAGCGGAG 325
 QY 98 ThrHisPheLeuThrAapGlyLeuArgAlaPheGluAlaArgValSerLygSerProPro 117
 DB 326 ACTCATTTCTTGGAGATGGAATGAGAGGCTTATGAAGCTGATGAGATTCTTAACACCT 385
 QY 118 GlnPheGlyPheSerAapAlaLeuLeuSerArgGlnLeuGlnAlaAlaCy 137
 DB 386 CAGTACGGGTTTACCCCTGATGCTTGTGATCAGCCGTAAGATTTCTCAAGCTGCTGT 445
 QY 138 GluThrLeuSerAapProArgSerArgArgGlyTyrAsnGlyLeuLeuAapAapGlu 157
 DB 446 GAACCTGATGCTAATGCTGCTTGAAGAGATTAATCAATGAGCATCGCTATGACCAT 505
 QY 158 GlnAlaThrValIleThrAapValProTrpAapLygValProGlyValLeuCyValleu 177
 DB 506 GCCGATACATCTCTCACTGATGTTCTTGGAGAACAGTCTTGGACATTTGCTGTG 565
 QY 178 GlnGluGlyGlyGluhrGluValleuValLeuArgValIGlyGlualeuLeuLySGlyArg 197
 DB 566 CAAAGAGCTGGGAGAGCGAGGAGTGTGAGAAATTTGGAGAGAAATCTTTGAAGAGAGA 635
 QY 198 LeuProLySerPheLySGlnAepValValLeuValMeAlaleuAlaPheLeuAapVal 217
 DB 626 CTTCCCAAGCTTTCAAGCAAGATGTTGCTTGGCGATGCGCATGCTGATGTCATAT 685
 QY 218 SerArgAapAlaMeAlaleuAapProProAapPheIleThrGlyTyrGlu 234

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2006, 12:44:39 / Search time 1193 Seconds
(without alignments)
4474.785 Million cell updates/sec

Title: US-10-600-070B-2

Perfect score: 4063

Sequence: 1 MEALSHVIGIGLSPFLCRPLP.....YEVFWKSGWKITGVTLAS 801

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWB.spool/US1060070/runat.21022006.112357.8602/app.query.fasta.1
-DB=N.Geneseg -QMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAHSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USRB=US1060070 @CGEN_1_1727@runat.21022006.112357.8602 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004ds: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4063	100.0	2406	12	ADJ38129	Adj38129 Arabidops
2	4051	99.7	2406	12	ADJ38208	Adj38208 Arabidops
3	4051	99.7	2637	12	ADJ38210	Adj38210 Placid d
4	4051	99.7	2679	13	ADT14901	Adt14901 Plant cdn

ALIGNMENTS

RESULT 1	ID	ADJ38129	standard; cDNA; 2406 BP.
XX	XX	ADJ38129;	
AC	ADJ38129;		
XX	XX	06-MAY-2004 (first entry)	
DT	XX		
XX	XX		
DE	XX	Arabidopsis thaliana Arc6-1 cDNA Seq1d.	
XX	XX		
KW	XX	prokaryotic type; placid division; Fun2; AR6; ARCS; Fzo; plant cell;	
KW	XX	agronomic; horticultural; crop plant; ornamental plant; woody plant;	
KW	XX	herbicide target; gene; ss.	
XX	XX		
OS	XX	Arabidopsis thaliana.	
XX	XX		
PN	XX	MO2004001003-A2.	
PD	XX		
PD	XX	31-DEC-2003.	
XX	XX		
PF	XX	20-JUN-2003; 2003MO-US019536.	
XX	XX		
PR	XX	20-JUN-2002; 2002US-0390140P.	
PR	XX	09-AUG-2002; 2002US-0402242P.	
XX	XX	20-JUN-2003; 2003US-00600070.	
XX	XX		
PA	XX	(UNMS) UNIV MICHIGAN STATE.	
XX	XX		

PI Oteryong KM, Vitcha S, Koksharova OA, Gao H;
XX WPI: 2004-082486/08.
DR P-PSDB; ADJ38202.

XX New isolated Ftn2, ARCS and/or Pzo-like nucleic acid sequences, useful
PT for further characterizing plasmid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.

XX Claim 1, SEQ ID NO 1, 287bp; English.

XX This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2406
Score:	4063.00	Matches:	801
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-600-070b-2 (1-801) x ADJ38129 (1-2406)

QY 1 MetGluAlaLeuSerHisValIGlyLeuSerProPheGlnLeuCyValGluPro 20
DB 1 ATGGAAGCTCTGAGTCACTGCGCATGCTCTCTCCCATTCCTCAATTATGCCGATTACCA 60
QY 21 ProAlaThrThrIleValArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
DB 61 CCGCGCAGCAAAAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCTCGCAGC 120
QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnThrSerThrIleCysSerAlaSer 60
DB 121 AATGGCCGACCTCTCTCTCGACTTCATTTCCATTCCTCCATTCCTCTCTCTCC 180
QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProPheSerIleAspArgPro 80
DB 181 TTGGCCACCCGCCACCAACCGCCACTCTGCTCTGCGCACCATCTATTTGATCCGCC 240
QY 81 GluArgHisValProIleProIleAspPheTrpGlnValLeuGlyAlaGlnThrHisPhe 100
DB 241 GAACGCCACGCTCCCATCCCATGATTCTTACCAAGATTTAGAGCTCAAAACCATTTTC 300
QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerLysProProGlnPheGly 120
DB 301 TTAAACCAATGGATCAAAAGCATTCGAAAGCTAGGTTTCGAAACCGCGCAATTCGCT 360
QY 121 PheSerAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
DB 361 TTCAGCGACGACGCTTAAATCAGCCGAGACAGATTTCTTCAAGCTGCTTCGAAACTCTG 420
QY 141 SerAsnProArgSerArgArgGluTrpArgGlnGlyLeuLeuAspArgGluAlaThr 160
DB 421 TCTAATCTCGGCTTAAAGAGAGTACAAATGAAGTCTTCTTGATGATGAAGAACTACA 480
QY 161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuCysValIleGlnGlnGly 180
DB 481 GTCACTCACTGATGTTCTTGGGATGAAGCTTCTGGGCTCTCTGTGATTCGAAAGAGT 540
QY 181 GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys 200
DB 541 GGTGAGACTGAGATGATTTCTTGGGTTGGTGAAGGCTCTGCTTAAAGAGAGTTGCTTAAG 600

QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
DB 601 TCGTTTAACCAAGATGTGGTTTATGATGCGCCTTGCCGTTCTCGATGTCTGAGAGGAT 660
QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTrpGluPheValGluGluAlaLeu 240
DB 661 GCATGTGGCAATGGATCCACTGATTTTATTTACTGCTTATGAGATGTTGTAAGAAAGCTTGG 720
QY 241 LysLeuLeuGlnGluGluGluValAspSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
DB 721 AAGCTTTTACAGAGAGAGAGAGCACTAGCTTTCGACCGGATTTTACGTGCAAAATTGAT 780
QY 261 GluThrLeuGlnGluIleThrProArgTrpValLeuGlnLeuLeuGlyLeuProLeuGly 280
DB 781 GAGACTTGGAAAGATCACTCCGCTTATGCTTGGACTACTTGGCTTACCGCTTGGT 840
QY 281 AspAspTrpAlaAlaValArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
DB 841 GATGATTACGCTGCGAAAGACTAAATGGTTTAAAGCGGTGCGGAATATTTGTGTGCT 900
QY 301 ValGlyGlyGlyValAlaSerAlaLeuValGlyLeuThrArgGluLysPheMetAsn 320
DB 901 GTTGAAGAGGTGAGACATCAGCTCTTGTGGGAGTTTGAACCGTAGAAGTTTATGAA 960
QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
DB 961 GAGCGCTTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTACCCCAAG 1020
QY 341 AsnIleProAlaGluSerPheGluValTrpGluValAlaLeuAlaLeuValAlaGlnAla 360
DB 1021 AATATTCACAGACAGTCAATTTGAAGTTTACGAAGTTGCACTGCTTGTGTGCTCAAGCT 1080
QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGln 380
DB 1081 TTATTTGTAAGAGCAACCTTTTACAGATCTGATAGCAATTCACGCAACTTCAG 1140
QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTrpAspThrArgAsnAsnTrp 400
DB 1141 CAGCTTAAGGTATGCTATGAGATTTCTCGGAGTGTGTATGATACAGCAATTAATTGG 1200
QY 401 GluIleAspPheGlyLeuGluTrpGlyLeuCysValaLeuLeuIleGlyLysValaArgGlu 420
DB 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTCTTATAGCAAAAGTTGATGAA 1260
QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlyTrpArgAsnProAlaIleVal 440
DB 1261 TGCCGTATGTGTGGCTTACAGAGATGATTCACAAATATGAAATCCAGCTATTTGTG 1320
QY 441 GluPheValLeuGlnAsnSerAsnArgAspAsnAspAspLeuProGlyLeuCysLys 460
DB 1321 GATTTGTTTGGAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGACATATGCAAA 1380
QY 461 LeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
DB 1381 TTGTTGAAAACTGGTGTGGAGGGGTGTCTTCTTCTAGGTTCGAGAACCAAAAGTTAA 1440
QY 481 LysPheLysLeuGlyLysAspTrpTrpAspAspProMetValLeuSerTrpLeuGluArgVal 500
DB 1441 AAATTTAACTCGGGACCTACATATGATGATCCATATGTTTGTACTTGGAAAGAGTG 1500
QY 501 GluValValGlnGlySerProLeuAlaAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
DB 1501 GAGGTAGTTCAAGGTTCTCTTATGCTGTGCTGCGCAACTATGCAAGATGGAGCGAG 1560
QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTrpThrAsp 540
DB 1561 CATGTAAAGCTAGTACTATGACGACCTGCAAGAAAGTTTCTTCCCTTACCATGAT 1620
QY 541 ArgAsnSerAlaGluProLysAspValGlnGluValPheSerValAspProValGly 560
DB 1621 AGAACTCGGCTGAAACCAAGATGTGCAAGACAGTGTATTAGTATGATCTGTGTGT 1680

QY 561 AsnAnValGlyATGAPGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1681 AACATATGAGCCGCTGATGCTGAGCCTGCTCTTTATGACGAAGCTGTAAGCCTCT 1740
 QY 581 GluAnPheGluThrAsnAPTyAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1741 GAAAACTTGGAAATGATGATTATGCAATTCGAGCTGGGCTCTCAGAGATGAGCTTGAT 1800
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuysGluAlaSerValIleLeu 620
 DB 1801 GAAACTGCTGTTGAAAGTCCGTTGCTGATATGTTAAAGGAGGCAAGTGAATCCCTA 1860
 QY 621 AlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGlnIleTyPheLeuIleSer 640
 DB 1861 GCTGCTGCTGGGCAATTTGACATTTCACTGCTTACGCCAGAAATTTCTTTAAAGC 1920
 QY 641 SerSerSerPheGlnArgIleAspMetValSerSerMetGluSerAspValAlaThrIle 660
 DB 1921 AGCTCATCTTTTCAACCAAGATATGCTTCTTATGAAATCTGATGCTGACATATA 1980
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 1981 GGGTCAGTCAAGCTGAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGACAG 2040
 QY 681 AsnIleValSerIleTyPheGlnIleIleSerSerLeuAlaPheGlyProAspHisArgIle 700
 DB 2041 AATATATATTCGAAGTGGCAGAAATTAAGTCTGCTTTGGGCTTGTACCGCAT 2100
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuIleTyPheThrAspArgAlaAla 720
 DB 2101 GAAATGTTACCAAGGTTTGGATGGGCGAAATGCTGAAGATTTGGACTGACAGAGCAGCT 2160
 QY 721 GluThrAlaGlnLeuGlyLeuValIleTyAspTyThrIleLeuIleSerValAspSer 740
 DB 2161 GAAATCGCGAGCTGGTGGTGGTTATGATTAATCACTGTTGAAACATCTGTTGACAGCT 2220
 QY 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluSerAla 760
 DB 2221 GTGACAGTCTCAGAGATGAAACCCGCTGCTGTGAGAAAGCACTCTGGAGAGCTGCT 2280
 QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyThrThr 780
 DB 2281 TGCTATCTGATTTGGTTCAATCCAGAAACAAATGCTGATGTCAGAACTTACACAA 2340
 QY 781 ArgTyArgIleValPheTyPheSerIleSerGlyTyPheIleThrGluGlySerValLeuAla 800
 DB 2341 AGATACGAAGTCTGCTGCTCAGAGTCAAGGAGGAAATCACTGAAGGCTCTGTTCTTGA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403
 RESULT 2
 ID ADJ38208 standard; DNA; 2406 BP.
 AC ADJ38208;
 DT 06-MAY-2004 (first entry)
 XX
 DE Plasmid division-related Arc6 orthologue gene 2.
 KW prokaryotic type; plasmid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 XX herbicide target; gene; de.
 OS Arabidopsis thaliana.
 PN WO2004001003-A2.
 XX 31-DEC-2003.
 PD 20-JUN-2003; 2003WO-US019536.
 PF

XX 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 XX (UNMS) UNIV MICHIGAN STATE.
 PI Oseeryoung KW, Vicha S, Koksharova OA, Gao H;
 DR WPI; 2004-082466/08.
 DR P-PSDB; ADJ38207.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plasmid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.
 PS
 CC This invention relates to novel prokaryotic type or plasmid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterising plasmid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.
 SQ Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 0 Length: 2406
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-600-070b-2 (1-801) x ADJ38208 (1-2406)
 QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 DB 1 ATGGAAGCTTGAGTACGTCGGCATTTGCTCTCCCATTTACATTATGCCATTACCA 60
 QY 21 ProAlaThrThrIleValLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40
 DB 61 CCGGCGACGACAAAGCTCCGACGTAGCCCAACACCTCTCAACTATCTCTCCGACG 120
 QY 41 LysThrAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 DB 121 AAATGGGCGACGCTCTCTCTCGACTTAATTTCACTCCGATTCCTCTCTCTCTCC 180
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 181 TTGCGCACCGCCACACACGCGCATCTGCTCTCCGACCATTTATGATTCGCC 240
 QY 81 GluArgHisValProIleProIleAspPheTyGlnValLeuGlyAlaGlnThrHisPhe 100
 DB 241 GAACGCCACGTCCTCCCATCCCATTTGATTTTACACAGGATTAAGAGCTCAAAACATTC 300
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleProProGlnIlePheGly 120
 DB 301 TTAAACCATGAAATTCAGAAAGCATTTGAAAGGTTTCGAAACCGCGCAATTCGGT 360
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCysGluThrLeu 140
 DB 361 TTCAGCGACGACCGCTTTAATCAACCGGAGACATTTCTCAACTGCTTCCGAAACTCTG 420
 QY 141 SerAsnProArgSerArgArgGlyTyArgGlnGlyIleLeuLeuAspAspGluGluAlaThr 160
 DB 421 TCTAATCTCTGCTCTGAAAGAGATCAATGAAGGCTCTTCTGATGATGAAGAGCTACA 480

QY 161 ValIlethraspValProTTPAspLeuValProGIValaleuCyValleuGluGly 180
 DB 481 GTATCATCTGATGTTCTTGGGATPAGGTTCTGTCCTCTCTGTGATTCAGAAAGGT 540
 QY 181 GlyIuThrcIuIleValleuArgValGlyIuAlaleuLeuIysGluArgProLys 200
 DB 541 GGTAGAGCTAGATGATGTTCTTGGGTTGGTGGAGGCTCTGCTTAAGAGAGGTTGCTTAC 600
 QY 201 SerPheIysGluIleValleuValMetAlaleuAlaPheIleuAspValSerArgAsp 220
 DB 601 TCGTTTAAAGCAAGATGTTGTTAGTTAGTGGCGCTTCGTTTCGATGTCGAGGGAT 660
 QY 221 AlMetAlaleuAspProProAspPheIleThrcIyTyrgIuPheValGluGluAlaLeu 240
 DB 661 GCTATGGCATTTGATCCACTGATTTTAACTGTTATGAGTTTGTGAGGAAGCTTGG 720
 QY 241 LysLeuLeuGluGluGluIleValleuSerSerIleuAlaProAspLeuArgAlaGlnIleAsp 260
 DB 721 AAGCTTTTACAGAGAGAGAGAGCAAGTACCTTGCAAGGATTTACGTCGCAAAATTTGAT 780
 QY 261 GluThrcIuGluIleThrcProArgTyValleuGluIleuGluIleuProLeuGly 280
 DB 781 GAACCTTTGAGAGAGATCACTCCGTTATGTTCTTGAAGCTTACCTTACCGCTTGGT 840
 QY 281 AspAspTyraIalaIyAsArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 DB 841 GATGATTAACCTCGGAAAGCACTAAATGTTTAAAGCGTGTGCGGAATATTTTGTGTCCT 900
 QY 301 ValGlyGlyGlyIyAlaSerAlaLeuValGlyIyLeuThrcArgIuLysPheMetAsn 320
 DB 901 GTTGAGAGAGGTGAGATCAGCTCTTGTGGGGTTTGACCCCTGAGAAATTTATGAT 960
 QY 321 GluAlaIleuLeuAspMetThrAlaAlaGluGluValIlePheLeuPheValAlaThrProSer 340
 DB 961 GAGCGGTTTTTACAAATGACAGCTGCTGAGAGAGGTATCTTTTGTAGCTAACCCCAAGC 1020
 QY 341 AsnIleProAlaIleuSerPheGluValTyrgIuValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1021 AATATTCACAGAGAGCTATTTGAAGTTTACAAATTCATCTGCTCTTGTGGCTCAAGCT 1080
 QY 361 PheIleGlyIyLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnIleuGln 380
 DB 1081 TTTATTTGTAAGAGCACACACCTTTTACAGATGCTGATTAAGCAATTCACAGCACTTCAAG 1140
 QY 381 GluAlaIyValMetAlaMetGluIleProAlaMetLeuTyraPheTrpArgAspAspTrp 400
 DB 1141 CAGGCTTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACAGGAATTAATGG 1200
 QY 401 GluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
 DB 1201 GAGATTAACCTTCGCTCTAGAAAGGAGCTCTGTGCACTGCTTATAGCCAAAGTTGATGAA 1260
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyraGlnProAlaIleVal 440
 DB 1261 TGCCTATGCTGTTGGCTTACAGATGAGATTCACAAATTAAGAAATCCAGCTATTTGTG 1320
 QY 441 GluPheValleuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
 DB 1321 GAGTTTGTGTTGGAAATTCAAATCGATGACATATGATATGATCCCTGAGCTATGCAAA 1380
 QY 461 LeuLeuGluThrcTrpLeuAlaGlyValValPheProArgPheAspAspThrLysAspLys 480
 DB 1381 TTGTTGAAACCTGTTGGCGAGGTTGTCTTTCTTAGGTTCAAGACACCAAGATTA 1440
 QY 481 LysPheLysLeuGlyAspTyraPheAspProMetValleuSerTyraLeuGluArgVal 500
 DB 1441 AAATTTAACTCGGAGACTACTATGATGATCTTATGTTTGAATTAAGTTGAAAGAGTG 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaIleThrMetAlaArgIleGlyValGlu 520
 DB 1501 GAGGTAGTTCAAGGTTCTCTTCTTATGCTGCTCAAGCTTAAGGCAAGATTTGAGCCGAG 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyraThrasp 540

DB 1561 CATGTAAAGCTAGTACTATGACAGGCACTGCAGAAAGTTTCTTCCCGCTTACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrcValPheSerValAspProValGly 560
 DB 1621 AGAACTCGGCTAACCCCAAGATGTGCAGAGACAGCTGTTTAGTGAATCTGTTGGT 1680
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1681 AACATGTAGGCGGTAGTGTGAGCTGGGTCTTTATTCAGAAAGCTGTAAGACCTCT 1740
 QY 581 GluAsnPheGluThrcAsnAspTyraIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1741 GAAACTTTGAAACTAATGATTTATGCAATTCGAGCTGGGTCTCAAGAGACTAGCTTGTAT 1800
 QY 601 GluThrcValGluMetSerValAlaAspMetLeuIleGluAlaSerValLysIleLeu 620
 DB 1801 GAACCTACTGTTAAATATGTCGTTGCTGATATTTAAAGAGCGCAAGTGTGAAGATCTTA 1860
 QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyraPheLeuLysSer 640
 DB 1861 GCTGCTGTGTGGCAATTTGACATTTTACCTGTCAGCCAGAAATATTTCTTAAAGC 1920
 QY 641 SerSerSerPheGluArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
 DB 1921 AGCTCATCTTTTCAAGCGCAAGATGATGTTTCTTATGTAATCTGATGTCGTACCATTA 1980
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgTrpAlaGlu 680
 DB 1981 GGGTCAGTCAAGCTGACGATTCAGATTCAGAAAGCACTTCCAGATGAGATGCTAGACCTGCAAG 2040
 QY 681 AsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
 DB 2041 AATATGATTCAGAAAGTGGCAAGATTAATGATCTCTGGCTTTGGGCTGATTCACCGCATTA 2100
 QY 701 GluMetLeuProGluValleuAspGlyArgMetLeuLysIleTrpThrAspArgAlaAla 720
 DB 2101 GAAATGTTACCAAGAGGTTTGGATGGCGCAATCTGAAGATTTTGGACTGACAGAGCAGCT 2160
 QY 721 GluThrcAlaGluLeuGlyLeuValTyraAspTyraThrcLeuLeuLysLeuSerValAspSer 740
 DB 2161 GAACCTCGGCAAGCTGTGGGTGTTTATGATTAATACCTGTTGAAACATATCTGTTGACAGCT 2220
 QY 741 ValThrcValSerAlaAspGlyThrcArgAlaLeuValGluAlaThrcLeuGluIleuSerAla 760
 DB 2221 GTGACAGCTTCACAGATGAGAACCCGTCCTGTGTGAACCACTCGAGAGTCTGCT 2280
 QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyraThrThr 780
 DB 2281 TGTCTATCTGATTTGTTTCATCCAGAAACAAATGCTATGATGTCAGAACCTTACACAAACA 2340
 QY 781 ArgTyraGluValPheTrpSerLysSerGlyTrpLysIleThrcGluGlySerValleuAla 800
 DB 2341 AGATTAACAAATTTTCTGCTCAAGTCAAGGTGGAAGAAATCACTGAAGGCTCTGTTCTTGCA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403

RESULT 3
 ADJ38210
 ID ADJ38210 standard; DNA; 2637 BP.

AC ADJ38210;
 DT 06-MAY-2004 (first entry)
 DE Plastid division-related Arc6 orthologue gene 3.

XX prokaryotic type; plastid division; Ftr2; ARG6; ARCS; Fzo; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; gene; ds.

OS Arabidopsis thaliana.
 XX PN W02004001.0003-A2.
 XX PD 31-DEC-2003.
 XX PF 20-JUN-2003; 2003MO-US019536.
 XX PR 20-JUN-2002; 2002JUS-0390140P.
 XX PR 09-AUG-2002; 2002JUS-0402242P.
 XX PR 20-JUN-2003; 2003JUS-00600070.
 XX PA (UNMS) UNIV MICHIGAN STATE.
 XX PI Oosteryoung KM, Vitsha S, Koksharova OA, Gao H;
 XX DR WPI; 2004-082486/08.
 XX DR P-PSDB; ADJ38209.
 XX PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 PS Disclosure; Fig 8; 287pp; English.
 XX
 XX CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (AR66), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.
 XX
 XX SQ Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 2637
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-600-070b-2 (1-801) x ADJ38210 (1-2637)
 QY 1 MetGluAlaLeuSerHisValGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 DB 114 ATGGAGCTCTGAGTCACTGGGCAATGGTCTCTCCCATTCCTCAATTATCCGATTACCA 173
 QY 21 ProAlaThrThyLeuArgArgSerHisAsnThrSerThrXlleCysSerAlaSer 40
 DB 174 CCGGCGACGACAAAGCTCCGACGTAGGCACACACCTCTCAACTATCTCTCGCGACG 233
 QY 41 LysTrpAlaAspArgLeuLeuSerArgPheAsnPheThrSerArgSerSerSerSer 60
 DB 234 AATATGGCCGACGCTCTCTCTCCGACTTCAATTTCACCTCCGATTCTCTCTCTCTCC 293
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 294 TTGCGCACCGCCACACACACGCGCACTCTGCTCTCCGACACATCTATTGATCGTCC 353
 QY 81 GluArgHisValProIleProIleAspPheTyrGlnValIleuGlyAlaGlnThrHisPhe 100
 DB 354 GAACGCGACGCTCCCATCCCATTTGATTCTACAGGTAATTAGAGCTCAACAACATTTCC 413
 QY 101 LeuThrAspGlyIleArgArgAlaPheGlnAlaArgValSerIleProPheGlnPheGly 120
 DB 414 TTAAACGATGGAATCGAAGAGCATTCGAAGCTTAGGGTTTTCGAAACCGCGCAATTCGGT 473
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCySGluThrLeu 140

DB 474 TTCAGGACGACGACCTTTAACTACCCGAGACAGATTCTTCAACTGTTCCGAACCTCG 533
 QY 141 SerAsnProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspArgGlnAlaThr 160
 DB 534 TCTAATCTCTCGCTCTGAGAGAGATCAATGAAGCTCTCTGATGATGAAGAGCTTACA 593
 QY 161 ValIleThrAspValProTrpAspLysValProGlyValLeuCyValIleuGlnGlnGly 180
 DB 594 GTGATCACTGATGTTCTTGGAATGAAGTTCCTGGGCTCTCTGATGATGAAGAGT 653
 QY 181 GlyIleThrGlnIleValLeuArgValGlyGlyAlaLeuLeuIleuGlyArgLeuProLys 200
 DB 654 GGTGAGACTGAGATGATGTTCTTCCGGTTGGTAGAGGCTCTGCTTAAGAGAGGTTCCCTAAG 713
 QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 714 TCGTTTAAGCAAGATGGTTTATGATAGCCCTTCTCGATGCTCGAGGAGT 773
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGlnAlaLeu 240
 DB 774 GCTATGGCATGGATGATCCATGATTTATTAAGCTGTTATGATGTTGTTGAGAAAGCTTGG 833
 QY 241 LysLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
 DB 834 AACCTTTTACGAG 893
 QY 261 GluThrLeuGlnGlnIleThrProArgTyrValLeuGlnLeuLeuGlnLeuProLeuGly 280
 DB 894 GAGACTTTGGAAGAGATCACTCCGCTTATGTTCTGGAAGCTTACCTGAGCTTACCTGAGT 953
 QY 281 AsnAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 DB 954 GATGATTTAGCGCTCGGAAGAGCTTAATGTTTAAGCGGTGTGGAGATATTTGTGCTCT 1013
 QY 301 ValGlyGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlyLysPheMetAsn 320
 DB 1014 GTTGAAGAGAGTGGAGACATCACTCTGTTGGGGGTTTGAACCGTGAAGAGTTATGAAT 1073
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
 DB 1074 GAGCGCTTTTACGATTAAGATGACAGCTGCTGAGAGAGTGAATCTTTTGTAGCTTACCCCAAGC 1133
 QY 341 AsnIleProAlaGlnSerPheGlnValTyrGlnValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1134 AATATTCACGACAGATCATTTGAAGTTTACGAAGTTGCACTTCTGTGGCTCAAGCT 1193
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1194 TTTATTTGTTAAGAGCACACCTTTTACAGATGCTGATAGCAATTCACGAACTTCAG 1253
 QY 381 GlnAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnAsnTrp 400
 DB 1254 CAGGCTTAAGGTAATGCTATGAGATTTCTGCGCATGTTGATGATACACGAAATTAATGG 1313
 QY 401 GlnIleAspPheGlyLeuGlnArgGlyLeuCyValAlaLeuLeuIleGlyLysValAspGln 420
 DB 1314 GAGATGACCTTCGCTTAAGAAAGGAGCTCTGCACTGCTTATAGCAAAAGTTGATGA 1373
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGlyAspSerGlyTyrArgAsnProAlaIleVal 440
 DB 1374 TGCCGTAATGTTGGGCTTAAACAGTGAAGATTCACATTAAGAAATCCAGCTATTTGTC 1433
 QY 441 GluPheValLeuGlnLysSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
 DB 1434 GAGTTTGTGTTGGAGAAATCAATCGTATGACAAAGATGATCTCCCTGAGATATCCAA 1493
 QY 461 LeuLeuGlnThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
 DB 1494 TTGTTGAAACCGGTGTCAGGGGTGTCTTCTTCAAGGTTCAGAACACCAAAAGTAA 1553
 QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGlnAlaGlyVal 500

Alignment Scores:

Pred. No.: 0 Length: 2679
 Score: 4051.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 13 Gaps: 0

US-10-600-070b-2 (1-801) x ADT14901 (1-2679)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
 DB 115 ATGGAACCTCGAAGTCACGTCGGCATTGCTCTCCCATTCACATTATATGCGATTACCA 174
 QY 21 ProAlaThrThrIleuArgArgSerHisAsnThrSerThrTrpIleCySerAlaSer 40
 DB 175 CCGCGCAGCAACAAAGCTCCGACGTAGCCAAACACTTACAACTATCTGCTCCGACAGC 234
 QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 DB 235 AATGGGCGCACCTCTTCTCTCCGACTTCACATTTCACCTCCGATTCTCTCTCTCTCC 294
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 DB 295 TTGGCCACCGCCACCAACCGCCACTCTGCTCTCCGCAACCATGATTGATGCTCC 354
 QY 81 GluArgHisValProIleProIleAspPheTrpGlnValLeuGlyAlaGlnThrHisPhe 100
 DB 355 GAACGCCACGTCCTCCATCCCATGATTTCTACCAAGATATGAGAGCTCAAAACACATTTC 414
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleProProGlnPheGly 120
 DB 415 TTAAACCATGGATTCAGAAAGACATTCAGAGCTAGGGTTTGAAACCGCCCAATTCGCT 474
 QY 121 PheSerAspAlaLeuIleSerArgGlnIleLeuGlnAlaAlaCySerGluThrLeu 140
 DB 475 TTGAGGACGACGCTTTATCATGCCGAGACAGATTCTTCAAGCTGGCTGGAAACTCTG 534
 QY 141 SerAsnProArgSerArgArgGluTrpAsnGluGlyLeuLeuAspAspGluAlaThr 160
 DB 535 TCTAATCTCGCTGTACAGAGAGATCAATCAATGAGCTCTTGTGATGATGAAGAGCTACA 594
 QY 161 ValIleThrAspValProTrpAspIleValProGlyAlaLeuCySerValLeuGlnGluGly 180
 DB 595 GTCTATCATGATGCTCTGGGATGAAGTCTCTGCTGTGTATGTCAGAAAGCT 654
 QY 181 GlyIleThrGluIleValLeuArgValGlyValAlaLeuLeuLeuGlyLeuArgProLys 200
 DB 655 GGTGAGACTGAGATAGTCTTCGGGTTGGTGAAGCTCTGCTTAAGAGAGGTTGCCCTAG 714
 QY 201 SerPheIleGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 DB 715 TCGTTTAAAGCAAGATGTGGTTTGTATGCGCTGCTTCGATGCTCGAGAGGAT 774
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTrpGluPheValGluGluAlaLeu 240
 DB 775 GCTATGCACTTGAATTCACCTGATTTTAACTGTTATGATGATTTGTTGAGAAAGCTTGG 834
 QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
 DB 835 AAGCTTTTACAGAGAGAAAGAGCAAGTACCTTCACCGGATTTACGTGCAAAATTCAT 894
 QY 261 GluThrLeuGluGluIleThrProArgTrpValLeuGluLeuLeuGlyLeuProLeuGly 280
 DB 895 GAGACTTTGGAAGAGATCACTCCCGCTTATGCTTGTGAGGCTACCTGCTTACCCCTTGGT 954
 QY 281 AspAspTrpAlaAlaLysArgLeuAsnGlyLeuSerGlyValAlaGlyAsnIleLeuTrpSer 300
 DB 955 GATGATTAACGCTGCGAAAGACATAAATGGTTTAAAGCGGTGGCGAAATATTTTGTGCT 1014
 QY 301 ValGlyValGlyAlaSerAlaLeuValGlyValLeuThrArgGluLysPheMetAsn 320
 DB 1015 GTTGGAGAGGTGAGATCACTCTTGTGGGGGTTTGAACCCCTGAGAACTTATGAT 1074

QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
 DB 1075 GAGCGGTTTTTACCAAGACAGCTGCGAGAGGTTGATCTTTTGTAGTACCCCAAGC 1134
 QY 341 AsnIleProAlaGlnSerPheGluValTrpGluValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1135 AATATTCAGAGAGCTATTGAAAGTTTGAAGATTCACCTGCTTGTGGCTCAAGCT 1194
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1195 TTTATTTGTAAAGAACCAACCTTTTACAGATGCTCATATGCAATTCAGCACTTCAG 1254
 QY 381 GluAlaLysValMetAlaMetGluIleProAlaMetLeuTrpAspThrArgAsnAsnTrp 400
 DB 1255 CAGGCTAAGGTAATGAGCTATGAGATTCCTGCGATGTTGTATGATACCGGAATTAATTGG 1314
 QY 401 GluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
 DB 1315 GAGATAGACTTCGGCTTAGAAAGGAGACTCTGTGCATCTGCTTATAGCAAAAGTTGATGAA 1374
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTrpArgAsnProAlaIleVal 440
 DB 1375 TGCCGTATGTGGTTGGCTTAGACAGTGAAGATTCACAATATAGAAATCCAGCTATTGTG 1434
 QY 441 GluPheValLeuGluAsnSerAsnArgAspAspAspAspLeuProGlyLeuCyLys 460
 DB 1435 GAGTTTGTGGAGAAATTCAAATCTGAAGACATATATATCTCCCTGACATATGCAAA 1494
 QY 461 LeuLeuGlnThrTrpLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
 DB 1495 TTGTTGAAACCTGGTTGGCAGAGGTTGCTCTTCTAGTTTCAGAGACCAAAAGATTA 1554
 QY 481 LysPheLysLeuGlyAspTrpTrpAspAspProMetValLeuSerTrpLeuGluArgVal 500
 DB 1555 AAATTTAACTCGGGGACTACTATGATGATCCATGTTTGAATCTTGGAAAGAGTG 1614
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
 DB 1615 GAGGTAGTCAAGGTTCTCTTATGCTGCTGCTGACCTATGGCAAGATTTGGAGCGAG 1674
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTrpThrAsp 540
 DB 1675 CATGTGAAGCTATGCTATATGACGACCTGCAAGAAATTTTCTTCCGCTATACAGAT 1734
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
 DB 1735 AGAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTATAGTATCCTGTTGGT 1794
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1795 AACATGTAGCCGTGATGTGTAGCCGTGGTCTTATTCAGAAAGCTGTAAGACCTCT 1854
 QY 581 GluAsnPheGluThrAsnAspTrpAlaIleArgAlaGlyValSerGlnSerSerValAsp 600
 DB 1855 GAATACTTTAATACTAATATGATTTATGCAATTCAGACTGGGCTTCAGAGAGTAGCTTGAT 1914
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu 620
 DB 1915 GAAACTACTGTTGAAATAGTCGTTGCTGATATGTAAAGAGGCAAGTGTGAAGATCTTA 1974
 QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTrpPheLeuLysSer 640
 DB 1975 GCTGCTGGTGTGGCAATTTGACATTCATCTGTCAGCAAGATTTTCTTAAAGC 2034
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAspValAlaThrIle 660
 DB 2035 AGCTCATCTTTTCAACGCAAGATATGCTTTCTTATGAAATCTAATGTCGCTACCATTA 2094
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 2095 GGGTCAGTCAAGCTGACATTCAGAAAGCACTTCCGAGATGATGCTGACAGCTGCAGAG 2154

QY 681 Asn11eValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
| | | | |
DB 2155 AATATAGTATCCAGAGTGGAGAGATTAACTCTGGCTTTGGGCTGATCACCGCAT 2214
| | | | |
QY 701 G1MeLeuProG1ValLeuAspGlyArgMetLeuLysIleTrpThrAspAlaAla 720
| | | | |
DB 2215 GAATGTATACAGAGGTTTGGATGGGCGAATGCTGAAGATTGGACTGACAGAGCGCT 2274
| | | | |
QY 721 GluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuLysSerValAspSer 740
| | | | |
DB 2275 GAAACCTCCGACGCTGGGTTGGTTTATGATTAATACCTGTTGAACCTATCTGTTGACGT 2334
| | | | |
QY 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGlnGlySerAla 760
| | | | |
DB 2335 GTGACAGCTTCAGAGATGGAAACCGTCTGCTGGTGAAGCACTCTGAGAGAGTCTGCT 2394
| | | | |
QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
| | | | |
DB 2395 TGTCTATCTGATTTGGTTCATCCGAAACCAATCTCTGATGTCAGAACCTACACACA 2454
| | | | |
QY 781 ArgTyrGluValPheTrpSerLysSerGlyTrpLysIleThrGlnGlySerValLeuAla 800
| | | | |
DB 2455 AGATACCAAGTTTCTGCTCCAGTCCAGGTTGGAAATCACTGAAGCTCTGTTCTTGCA 2514
| | | | |
QY 801 Ser 801
| | |
DB 2515 TCA 2517
| | |
RESULT 5
ID ADJ38135 standard; cDNA; 2406 BP.
XX ADJ38135;
AC ADJ38135;
DT 06-MAY-2004 (first entry)
XX Arabidopsis thaliana AtFtn2 cDNA SegID9.
DE
XX prokaryotic type; plasmid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX Arabidopsis thaliana.
OS
XX
PN MO2004001003-A2.
XX
XX 31-DEC-2003.
PD
XX 20-JUN-2003; 2003MO-US019536.
PF
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-0060070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
PA
XX
PI Osteeryoung KM, Vitha S, Kokscharova OA, Gao H;
DR WPI; 2004-082486/08.
DR P-PSDB; ADJ38203.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plasmid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX
PS Claim 1; SEQ ID NO 9; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary

CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2406
Score:	4049.00	Matches:	799
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	2
Query Match:	99.7%	Indels:	0
DB:	12	Gaps:	0

US-10-600-070b-2 (1-801) x ADJ38135 (1-2406)

QY	1	MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuGlyArgLeuPro	20
DB	1	ATGGAAGCTCTGATGTCACGTCGGCATTTGGTCTCTCCCATTCATTAATGATGCGATTACCA	60
QY	21	ProAlaThrThrLysLeuArgArgSerHisAsnThrSerThrThrIleCysSerAlaSer	40
DB	61	CCGGCAGACGACAAAGCTCCGACCTAGCCACACACCTTACAACTATCTGCTCCGACG	120
QY	41	LysTrpAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer	60
DB	121	AAATGGGCGGACCGCTCTGCTCCGACCTTCAATTTACCTCCGATTCTCTCTCCCTCC	180
QY	61	PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro	80
DB	181	TTGCGCACCGCCACACACCGCACCTCTGCTCTCTGCGACCACTATGATCGTCCC	240
QY	81	GluArgHisValProIleProIleAspPheTyrGlnValLeuGlnValAlaGlnThrHisPhe	100
DB	241	GAAAGCCAGCTCCCAATCCCATTAATTTCTACAGATTTAGAGCTCAAAACATTTTC	300
QY	101	LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly	120
DB	301	TTAAACGATGGATTCAGAGAGCATTCGAAGCTAAGGTTTGGAAACCGCGCAATTCGGT	360
QY	121	PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu	140
DB	361	TTCAGCGACGACCGCTTAAATCAAGCCGAGCAGACATTTCTCAAGCTCTGCGAAACTCG	420
QY	141	SerAsnProArgSerArgArgGluTyrAsnGlnGlyLeuLeuAspAspGlnGluAlaThr	160
DB	421	TCTAATCTCTGCTGTAAGAGAGTACAAAGAAAGTCTTCTGATGATGAAGAGCTTACA	480
QY	161	ValIleThrAspValProTrpAspLysValProGlyAlaLeuGlyValLeuGlnGlnGly	180
DB	481	GTCATCACTGATTTCTCTGGATTAAGTTCCCTGGGGCTCTCTGATTTGCAAGAAAGT	540
QY	181	GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluArgLeuProLys	200
DB	541	GGTGAAGCTGAGATGATGTTCTTCGGTTGGAGAGCTCTGTTAAGAGAGGTTGCTTAAG	600
QY	201	SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp	220
DB	601	TGCTTTAAGCAAGATGGTTTATGATAGCCGCTTCTCATGTCTCGAGGAGT	660
QY	221	AlMetAlaLeuAspProProAspPheIleThrGlyTyrGlnPheValGluGluAlaLeu	240
DB	661	GCTATGGCATTTGATTCACCTGATTTTATTAATGAGTTTGTGAGAGAGCTTTG	720
QY	241	LysLeuLeuGlnGlnGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp	260
DB	721	AACTTTTACAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTGCAAAATTTGAT	780
QY	261	GluThrLeuGlnGluIleThrProArgTyrValLeuGlnLeuLeuGlyLeuProLeuGly	280
DB	781	GAGACTTTGGAGAGATCACTCCGCGTATGTTCTTGAAGTACTTGGCTTACCGCTTGGT	840

QY 281 AspAspTyrAlaAlaIleValArgLeuAsnGlyLeuSerGlyValAlaGlnIleLeuTyrSer 300
 DB 841 GATGATTACCGCTCGGAAAGACTAAATGCTTTAAGCCGTGTGCGAAATATTTTGTGCTCT 900
 QY 301 ValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlyLeuPheMetAsn 320
 DB 901 GTTGAAGAGGTGAGCATGAGCTCTTGTGGGGGTTGACCCGTGAAGATTATGAAAT 960
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
 DB 961 GAGCGCTTTTATGAATGACAGCGCTGACAGCGGTTATCTTTTGTAGCTAACCCCAAGC 1020
 QY 341 AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAla 360
 DB 1021 AATATTCACAGAGTATTTGAGTTTACGAACTTGCATCTGTGTGCTCAAGCT 1080
 QY 361 PheIleGlyLysProAlaLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 DB 1081 TTTATTGCTAAGAACCAACACCTTTTACAGGATGCTATTAAGCAATTCAGCAACTTCAAG 1140
 QY 381 GluAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnMetTyr 400
 DB 1141 CAGGCTAAGGTAATGCGCTATGAGATTCCTGCGATGTTGATGATACAGCAATAAATTGG 1200
 QY 401 GluIleAspPheGlyLeuGluValArgGlyLeuCysAlaLeuLeuIleGlyLysValAspGlu 420
 DB 1201 GAATATGACTTCGGTCTAGAAAGGGGACTGTGTGCTGCTTATAGCAAAAGTTGAAGAA 1260
 QY 421 CysArgMetTyrPheGlyLeuAspSerGlyLysAspSerGlnTyrArgAsnProAlaIleVal 440
 DB 1261 TGCGGTATGTGGTGGGCTTAGACAGTGAAGATTCAAAATRAGGATTCACGCTATTTGCTG 1320
 QY 441 GluPheValLeuGluLysAsnSerAsnArgAspAsnAspAspLeuProGlyLeuCysLys 460
 DB 1321 GAGTTTGTGTTGAGATTCGAAATTCGTGATGACATATGATCTCCCTGAGCTATGCAAA 1380
 QY 461 LeuLeuGluThrTyrPheLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
 DB 1381 TTGTGGAAACCTGGTGTGGCGGGGTGTCTTCTTCTGATTCAGAGCAACCAAGATAA 1440
 QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGluLysVal 500
 DB 1441 AATTTTAACTCGGGGACTATGATGATGATCTTATGAGTTTGAAGTTACTTGAAGAGCTG 1500
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaAlaThrMetAlaArgGlyIleGlyAlaGlu 520
 DB 1501 GAGGTATGTCAGGCTTCTCTTATGCTGCTGCTGCACTTGGCAAGATTGAGCCGAG 1560
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
 DB 1561 CATGTGAAGCTAGTGTATGACAGGCACTGCAAGAAATTTTCTTCCCTATACAGAT 1620
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGlyThrValPheSerValAspProValGly 560
 DB 1621 AGAACTCGGCTGAACCAAGATGTGCAAGACAGATGTTTACTGTAGATCTCTTGTGCT 1680
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 DB 1681 AACAAATGTAAGCGGTGATGTGAGCTGTGTCTTATTCAGAAAGCTGTAAAGCCCTCT 1740
 QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 DB 1741 GAAAACTTGAACCTAATGATTAATGCAATTCGAGCTGCGGTCTTCAGAGATAGCGTTGAT 1800
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluLysSerValLysIleLeu 620
 DB 1801 GAATCTACTCTTGAATATGCTCGTGTGCTGATATGTAAAGAGCAAGCAATGTGAAGATCTTA 1860
 QY 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSer 640
 DB 1861 GCTGCTGTGTGCAATTTGAGCTGATTTCACTGTTCAAGCAAGATTTTCTTAAAGC 1920
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660

DB 1921 AGCTCATCTTTTGAACGCAAGATATGCTTTCTTATGATGATCTGATGCTGACATA 1980
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 DB 1981 GGGTCACTGACAGCTGACATTCAGTGAAGCACTTCCAGAAATGATGCTGAGACTGACAG 2040
 QY 681 AsnIleValSerLysTyrPheGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
 DB 2041 AATATGATATCCAAATGCGCAAGAGATTAAGTCTCTGCTTGTGGCTGATACCGGATA 2100
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAla 720
 DB 2101 GAATGTATACAGAGCTTTTGGATGCGCAATGCTGAAGATTTGACTGACAGAGCAAGCT 2160
 QY 721 GluThrAlaGluLeuGlyLeuValTyrAspTyrThrLeuLeuLysLeuSerValAspSer 740
 DB 2161 GAAACTGCGCAGCTGGGTGGTTTATGATTAACACTGTTGAATCTATCTGTTGACAGT 2220
 QY 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 760
 DB 2221 GTGACAGTCTCAGCAGATGGAACCCGTGCTGTGTGAGCAACTGTGAGAGAGCTGCT 2280
 QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
 DB 2281 TGTCTATCTCATTTGGTTCATCCGAAACAAATCTCTATGATGTCAGAACTTACACAA 2340
 QY 781 ArgTyrGlyValPheTyrSerLysSerGlyTyrIleLysIleThrGluGlySerValLeuAla 800
 DB 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGTGAGGTAATCACTGAAGGCTCTGTTCTTGA 2400
 QY 801 Ser 801
 DB 2401 TCA 2403
 DB 2401 TCA 2403
 RESULT 6
 ADJ38130
 ID ADJ38130 standard; DNA; 3667 BP.
 AC
 ADJ38130;
 DT 06-MAY-2004 (first entry)
 XX
 DE Arabidopsis thaliana Arc6-1 genomic DNA Segid2.
 XX
 KW prokaryotic type; plasmid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004001003-A2.
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI OsterYoung KM, Viltha S, Koksharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 DR P-PDB; ADJ38202.
 XX
 PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plasmid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX

Db 481 ATGGAAGCTCTGATCAGCGTCGCGATTCCTCTCCCATTCACATTAATGACCGATTACCA 540
 QY 21 ProAlaThrThrIleuArgArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
 Db 541 CCGGCACAGCAAAAGCTCCGACGTAGCCACACACCTCTACAACTACTGCTCCGCGACG 600
 QY 41 LysTrpAlaAspArgLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 Db 601 AATATGGGCGGACCGCTCTCTCTCCGACTTCATTTACCTCCGACTTCCTCTCTCTCC 660
 QY 61 PheAlaThrAlaThrThrThrAlaThrLysLeuValSerLeuProProSerIleAspArgPro 80
 Db 661 TTGGCCACCGCCGACCAACCGCCACTCTGCTCTCTGCGCACCATCTATTGATTCGTC 720
 QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlnValAsnThrHisIlePhe 100
 Db 721 GAACGCGACGTCCTCCATCCCATGATTCCTTACAGGATTTAGAGCTCAACACACTTC 780
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
 Db 781 TTAAACGATGGAATCCAGAGAGCAATTCGAAAGCTTAGGGTTTCGAAACCGCGCAATTCG 840
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGlnThrLeu 140
 Db 841 TTCAGGAGACGACTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTTCGAAACTCTG 900
 QY 141 SerAsnProArgSerArgArgGlyTyrAsnGlnGlyLeuLeuAspAspGlnGluAlaThr 160
 Db 901 TCTAACTCTCGCTCTGAAAGAGATCAATGAAGGCTCTTCTGATGATGAAGAGCTAC 960
 QY 161 ValIleThrAspValProTrpAsp----- 168
 Db 961 GTATCACTGATGTCTTGGAATAGGTAATTCGATTCGAAATTAATAAGTTCTTC 1020
 QY 169 -----LysValProGlyAl 173
 Db 1021 GTTTTAATTTCAAGAAATGGAATAAGGAAGAACTTTATCTAGTAAGGTTCTCGGGGC 1080
 QY 173 AlaCysValLeuGlnGlnGlyGlyGlnThrGlnIleValLeuArgValGlyGlnAla 193
 Db 1081 TCTCTGTGTATGTCAGAAAGGTGGTGAGACTGAGATGTTCTTCGGGTGGTGAAGGCTCT 1140
 QY 193 UleuLysGluArgLeuProLysSerPheLysGlnAspValValLeuValMetAlaLeuAl 213
 Db 1141 GCTTAAAGGAGAGGTGCTTAAGTGTCTTAAGCAAGATGTGGTTTAAGGCGCTTGC 1200
 QY 213 AspLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyr 233
 Db 1201 GTTTCGATGTCTCCAGGGATGCTATGGCATTTGGATCCACCTGATTTTATCTGGTTA 1260
 QY 233 GlnPheValGlnGlnAlaLeuLysLeuLeuGln----- 244
 Db 1261 TGAGTTGTTGAGGAAGCTTTGAAGCTTTTACAGGTAGTTGACTTCCTTGGTAATTG 1320
 QY 244 ----- 244
 Db 1321 ACAGCGCTTGGCTTTATAGAACTTTGATTTGATATCTTTGATTTAGCTTGTGTA 1380
 QY 245 -GlnGlnGlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGlnThrLeuGln 264
 Db 1381 GAGAGAAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTAAGAGCTTTGGA 1440
 QY 264 uGlnIleThrProArgTyrValLeuGlnLeuLeuGlnLysLeuProLeuGlnLysAspTyrAl 284
 Db 1441 AGAGATCACTCCGCTTATGTCTTGAGACTTGGCTTACCGCTTGTGTGATTTACGC 1500
 QY 284 aaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGly 304
 Db 1501 TGGGAAAAAGACTAATGTGTTAAGCGGTGTGCGGAATATTTGTGTGTGGAGAGG 1560
 QY 304 yGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlyLysPheMetAsnGlnAlaPheLe 324
 Db 1561 TGGAGCATCAAGCTCTTGTGGGGGTTTGACCCGTGAAGAGTTTATGAAATAGGCGTTT 1620

QY 324 uArgMetThrAlaAlaGlnGlnVal----- 332
 Db 1621 ATGAATGACAGCTGTGAGACAGT-ATACAGTTTAGATACCTTTTATTTCTTTAGC 1679
 QY 333 -----AspLeuPheVal 336
 Db 1680 ATGATATTAATTTAGGTTTCATATTTTAATGATGTGTGTGTAGGTTGATCTTTTGT 1739
 QY 336 lAlaThrProSerAsnIleProAlaGlySerPheGluValTyrGlnValAlaLeuAlaLe 356
 Db 1740 AGCTACCCCAAGCAATATTTCCAGAGATCATTTTGAAGTTTACGAAAGTTGACTTCCT 1799
 QY 356 uValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnP 376
 Db 1800 TGTGCTCAAGCTTTTATTTATGTAAGGCACACCTTTTACAGAGATGCTGATTAAGCAAT 1859
 QY 376 eGlnGlnLeuGlnGlnAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspTh 396
 Db 1860 CCAGCAACTTCAGCAGGCTTAAGGTAATGCTATGAGATTCCTCGCATGTTGTATGATAC 1919
 QY 396 rArgAsnAsnTrpGlnIleAspPheGlyLeuGlnArgGlyLeuCysValaLeuLeuIleG 416
 Db 1920 ACGAATATATGGAGATAGACTCGCTTGAAGAGGACTCTTGCACTGCTTATAGG 1979
 QY 416 yLysValAspGluCysArgMetTrpLeuGlyLeuAspSerGlnLysArgAs 436
 Db 1980 CAAGTGAATGAATGCCATGTGTGGCTTACAGACTGAGATTCACAAATATAGAA 2039
 QY 436 nProAlaIleValGlnPheValLeuGlnAsnSerAsnArgAspAspAspLeuPr 456
 Db 2040 TCCAGCTATTTGAGAGTTGTTTGGAGAAATTCATCTGATGTCMAATGATGATCTCC 2099
 QY 456 oGlyLeuCysLysLeuLeuGlnThrThrLeuAlaGlyValValPheProArgPheArgAs 476
 Db 2100 TGGACTATGCAATTTGTTGAAACCTGTGTGAGGGGTGTCTTCTTCAAGTTTCAGAA 2159
 QY 476 pThrLysAspLysLysPheLysLeuGlnLysAspTyrTyrAspAspProMetValLeuSerTyr 496
 Db 2160 CACCAAGATTAATAAATTTAACTCGGGGACTACATGATGATCCATGTGTTTGAGTTA 2219
 QY 496 rLeuGlnArgValGlnValValGlnGlySerProLeuAlaAlaAlaIleThrMetAlaAr 516
 Db 2220 CTGGAAAGGTGAGAGTATGTTACAGGGTCTCTCTTACGTGCTGCGACTATGCGAAG 2279
 QY 516 gIleGlyAlaGlnHisValValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSe 536
 Db 2280 GATTGAGCGGAGCATGTAAAGCTATGTCAGGCACTGCAAGAAAGTTTCTCTTC 2339
 QY 536 rArgTyrThrAspArgAsnSerAlaGlnProLysAspValGlnGlnThrValPheSerVal 556
 Db 2340 CCGCTATACAGATAGAAATCTCGGCTAACCAGAGATGTGCAAGAGACAGTGTATTAGTGT 2399
 QY 556 lAspProValGlyAsnAsnValGlyArgAspGlyGlnProGlyValPheIleAlaGlnAl 576
 Db 2400 AGATCTGTGTTGTPAACATGATGAGCGGTGATGCGCTGAGCGGTCTTATTTGCAAGAC 2459
 QY 576 aValArgProSerGlnAsnPheGlnThrAsnAspTyrAlaIleArgAlaGlyValaSerG 596
 Db 2460 TGTAAACCTCTGAAAACTTTGAAACTATATGTTTGAATTCAGAGCTGGGGTCTCAGA 2519
 QY 596 uSerSerValAspGlnThrThrValGlnMetSerValAlaAspMetLeuLysGlnAlaSe 616
 Db 2520 GAGTACCGTTGAGAAACTATCTGTAAATGTCCTGTCTGATATGTTAAAGAGCGCAAG 2579
 QY 616 rValLysIleLeuAlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyr 636
 Db 2580 TGTGAAGATCTTACGTGCTGTGCGCAATTTGACATTTCACTGTTCAGCCAGAAAGTA 2639
 QY 636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetLysSerAs 656
 Db 2640 TTTCTTTAAAGACGCTCATCTTTTCAACGCCAAGGATATGTGTTCTTCTATGGAATCTGA 2699

Oy	656	pval1a2rr1le	-----	660
Db	2700	TCGCGCTACACATAGGTATGATTAAATGATGCAATTTTCAATATATCTGCATTGCTCAAAAT	2755	
Oy	660	-----	660	
Db	2760	ATGCTTTGTTTGTGAGCTAAGAACATAGTTCCTTAATAATCATGTGCCAAAGTTGTAC	2819	
Oy	660	-----	660	
Db	2820	CAAGATTAAACAATTGCTGAGTAATAATTTCATTAATTAATGCTTGAAATTTTGTGATCAA	2879	
Oy	660	-----	660	
Db	2880	ACTGTAGACAGAAATGTAAATTTTCACTCTCAACATTTCTGTTTGAATTAACGTAGAGATTA	2939	
Oy	660	-----	660	
Db	2940	GAGATTGCGCTTAGTGTGGCTTTGTCCAACTTTCTTCTTGATTTTCTTTTCCATT	2999	
Oy	661	--GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG	680	
Db	3000	TAGGGTCACTGACGACCTGACGATTCAGAAAGCACTTCCAGATGGATCTGAGACTGACG	3059	
Oy	680	luAsnIleValSerIleTrrpGlnIlySerIleuAlaPheGlyProAspHisArgI	700	
Db	3060	AGAAATATAGTATCCAAAGTGGCAGAAAGATTAAAGTCTGTGGCTTTGGGCTGATCACCGCA	3119	
Oy	700	1eGluMetLeuProGlu-----	705	
Db	3120	TAGAAATGTTACACAGAGGTGAGGGAATTAATCTACAAATTCATCAATGTTGTGAAAAGT	3179	
Oy	706	-----	ValLeuAspGly	709
Db	3180	TTGGACATGATTATAGTCTGGTGCCTTGTGATTCTGTATTATTAAGTTTGGATGGG	3239	
Oy	710	ArgMetLeuIySileTrrpThrAspArgAlaAlaGluThrAlaGlnIleuGlyLeuValTyr	729	
Db	3240	CGAATGCTGAAGATTGTGACTGACAGACACTGAACTGAGCGCAGCTTGGGTGGTTGATTA	3299	
Oy	730	AspTyrThrLeuLeuIyLeuSerValAspSerValThrValSerAlaAspGlyThrArg	749	
Db	3300	GATTATACACTGTTGAAACCTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAAACCGGT	3359	
Oy	750	AlaIleuValGluAlaThrLeuGlnGluSerAlaCysLeuSerAspLeuValHisProGlu	769	
Db	3360	GCTCTGGTGGGAAGCACTCTGGAAGAGCTCTGCTGTATCTGAATTTGGTTCATCCAGAA	3419	
Oy	770	AsnAsnAlaThrAspValArgThrTyrThrThrArgTyrGluValPheTrrpSerLysSer	789	
Db	3420	AACAAATGCTACTGATGTGCAGAACCTACACAAACAAGATTCGAAGTTTCTGGTCCAAAGTCA	3479	
Oy	790	GlyTrrpIySileThrGlnGlySerValIleuAlaSer	801	
Db	3480	GGGTGGAAATACACTGAAAGCTCTGTCTTCTTCATCA	3515	
RESULT 8				
ADJ38206				
ID	ADJ38206 standard; DNA; 2283 BP.			
XX	ADJ38206;			
XX	06-MAY-2004 (first entry)			
XX	Plastid division-related Arc6 orthologue gene 1.			
XX	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;			
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;			
KW	herbicide target; gene; ds.			
XX	Oryza sativa.			
XX	MO2004001003-A2.			
PN				

XX 31-DEC-2003.
PD
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (DNMS) UNIV MICHIGAN STATE.
PI Oseeryoung KW, Vilcha S, Koksharova OA, Gao H;
XX
XX WPI: 2004-082486/08.
DR P-PSDB; ADJ38205.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8, 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (AR66), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX
XX Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	2.6e-154	Length:	2283
Score:	1775.50	Matches:	390
Percent Similarity:	62.5%	Conservative:	119
Best Local Similarity:	47.9%	Mismatches:	213
Query Match:	43.7%	Indels:	93
DB:	12	Gaps:	16
US-10-600-070B-2 (1-801) x ADJ38206 (1-2283)			
Qy	12 SerProPheGlnLeuCySarGluuPProProlaThrThnrlYsLeuArgSerHlAsn	31	
Db	40 GCGCCATTGGCTTTCCTCCCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCT	99	
Qy	32 ThnSerThnrlLeCyS---SerAlaSerlYlPAlaAspArgLeuLeuSerAspHe	50	
Db	100 CACCCCTCCGCTCCGCGCGCGCGGAGCGCTGGCGGAACGCTCTTGCGCGACTTC	150	
Qy	51 AsnPhe-----ThnSerAspSerSerSerSerPheAlaThnrlThr	65	
Db	160 CACCTCTCCGCCACGCGCGCGCTCCGACCGCGCGCTCCGCGCGCGCGCGCGCGCT	210	
Qy	66 ThnThrlAlaThnrlValSerleuPProProlSerlLeaAspArgPProGluArgHlValPro	85	
Db	220 GCGCGCTCCGCGCTCCCTTCGTCGCGCTTCCCGGACGCGCGCGGAAACGCTCCCTCCG	270	
Qy	86 lLeProIlleAspPheYrGlnValLeuGlYalaglThnrlAspHeLeuThnrlAspGlylLe	105	
Db	280 CTCGAAGTCGATTTCACAAAGTTCATAGGGGAGAGGCACATTCCTTGCGGATGGATCC	330	
Qy	106 ArgAlaAlaPheGlnAlaArgValSerlYrPProGlnPheGlyPheSerAspAspAla	125	
Db	340 AGAAGGCGCTTGAGGACGAGATACCAAGCACGCGCAGTGTGCTACAGACGAGATGCT	390	
Qy	126 lLeuIlleSerArgArgGlnlLeuGlnlAlaIaCyasGluPrlrLeuSerAsnProlArgSer	145	
Db	400 CTGTGTGTGTGTGTGACAAATGCTGGAGATGTGGCCATCATCTTCATGAAACCAACTCC	455	


```

Db      1 GCAGTTGCAATGCTGGGGNGAATTCACGTGAAATTTCAAGAACGAGCCTTCTTG 60
Qy      325 ArgMetThrAlaIaIaGluGlnValAspLeuPheValAlaThrProSerAsnIleProAla 344
Db      61 CATATGACTGACACTGAGCAGGTGATTTATTTGACTCTACCCCACTAATATCCCGGCA 120
Qy      345 GluSerPheGluValTyrGluValAlaIaLeuAlaLeuValAlaGlnAlaPheIleGlyLys 364
Db      121 GAAAGCTTTGAAGTTTATGCGGTGGCTCTTCCGCTTCTTCTCAAGCCTTTGTTGGTAA 180
Qy      365 LysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnGlnAlaLysVal 384
Db      181 AAACCTCATCACAATTCAGATGCTGAAACCACTTCCAGAAACCTTCAGCAGCTTAAGTGA 240
Qy      385 MetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnProGluIleAspPhe 404
Db      241 ACAGCTGTAGACATTCCTTTCACMCACTATATACCAAAAGAACGCTGAGATAGACTTT 300
Qy      405 GlyLeuGluArgGlyLeuGlyValAlaLeuLeuIleGlyLysValAspGluCySerMetTyr 424
Db      301 GCTTTGAGAGAGGAGACTCTTCTTCACTTCTTCAAGGAGCCTTGATGACAGTCCGTCGG 360
Qy      425 LeuGlyLeuAspSerGlnAspSerGlnTyrArgAsnProAlaIleValGluPheValLeu 444
Db      361 TTGGGCTTACAGCATATGATTCACATATAGAAATCCATCTGTTGTAGACTTTGCTTGG 420
Qy      445 GluAsnSer---AsnArgAspAspAspAsp-----AspLeuProGlyLeuGlyLys 460
Db      421 GAGAACTCAAAAGATGACATGACAAATGACAAATGACAAATGATCTTCTTGACCTTTGGCAG 480
Qy      461 LeuLeuGluIleThrTyrPheValAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
Db      481 CTATTGGAGACGTCGTGTTGATGAGAGCTGATTTCCCAAGCTTTAGAGACACCAAGACATA 540
Qy      481 LysPheLeuLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGluArgVal 500
Db      541 GAGTTGACGACTGGAGACTACTATGATGATCCCTACGCTTGAAGTACTTGAAGAGGCTG 600
Qy      501 GluValValGlnGlySerProLeuAlaAla 510
Db      601 GATGGCACTAATGTTCCACCTTACGCT 630

RESULT 11
ADJ38215 standard; cDNA; 660 BP.
ID ADJ38215 standard; cDNA; 660 BP.
XX
AC ADJ38215;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue cDNA 8.
XX
KW prokaryotic type; plastid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
OS Medicago truncatula.
XX
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
XX
PR 09-AUG-2002; 2002US-040242P.
XX
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Oeeryoung KW, Vitcha S, Koksharova OA, Gao H;
XX
DR WPI; 2004-082486/08.

```

```

XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287pp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6,5e-57 Length: 660
Score: 718.00 Matches: 141
Percent Similarity: 79.5% Conservative: 35
Best Local Similarity: 64.1% Mismatches: 34
Query Match: 17.7% Indels: 10
DB: Gaps: 4
XX
US-10-600-070B-2 (1-801) x ADJ38215 (1-660)
Qy      25 LysLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSerLysTyrAlaAsp 44
Db      15 AAACCTCAACGCTCCAT-----TCCTCGCGGCTCCGCCACAGTAAATGGCGGAG 68
Qy      45 ArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSerPheAlaThrAla 64
Db      69 CGACTCATTTCCGATTTCCCAATTCCTCGGCGACACTCTCTTCCCTCCACACACCACC 128
Qy      65 ThrThrThrAlaIleThrLeuVal---SerLeuProProSerIleAspArgProGluArgHis 83
Db      129 TCCGCCACAGTCCCTCTCACTCTCTTACCTCTTACCTCTCCGATA-----GAAACCCAC 179
Qy      84 ValProIleProIleAspPheTyrGlnValLeuGlyValAlaGlnThrHisPheLeuThrAsp 103
Db      180 GTGTCACTCCCTCTCGACCTGTCGAATCAATCTCGGCGGAAACGATTTCTCGGTGAT 239
Qy      104 GlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGlyPheSerAsp 123
Db      240 GGTATTGCGAGAGCTTATGAGGAAATCTCGAACCTCTCTCAGATGCTTCACTAAT 299
Qy      124 AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeuSerAsnPro 143
Db      300 GAAGCTTTGATTAAGTGTGCTGTCAGATCTTCAAGCTGTGTAACCTTACGTATCCT 359
Qy      144 ArgSerArgArgGluTyrAsnGlnGlyLeuLeuAsp-----AspGluGluAla 159
Db      360 GCTTCAAGAAAGAGATATATCAAAAGCTCTGTCAGATGAAGAAAGCAAGATGCAATCT 419
Qy      160 ThrValIleThrAspValProTyrAspLysValProGlyAlaLeuCyValLeuGlnGlu 179
Db      420 TCATTTCTCACTGAATATCCCTTTCGACAAAGTTCTTGAGACTTGTGCGGTGCAAGAA 479
Qy      180 GlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuArgGluArgLeuPro 199
Db      480 GCTGAGAGAGCGAGTTGGTGGCTTGCATTTGGAGGGGTTTATCTAGAGAGAGCTTACCG 539
Qy      200 LysSerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArg 219
Db      540 AAGATGTTAAGCAAAATGTGTGCTGCTATGCGCTTGCAATATTTGACGTTTCTAAG 539
Qy      220 AspAlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAla 239
Db      600 GATGCTATGGCTTTGTCCCGCCAGATTTCAATGTTGTTGAGATGCTGGAAGGCGCA 659

```



```

Db      586  GGTCAAGAG-----CTGCTAGTACGTGAAGCT-----TTG 615
Qy      252  AlaProAspLeuArgAlaGlnIleAspGluThrLeuGlnGluIleThrProArgTyrVal 271
Db      616  TTCTCCAGTTCACAGCGCAAAATTCAGCGTGAATTCACAAATTCGCGCCATTCGAAT 675
Qy      272  LeuGlnLeuLeuGlnLeuProLeuGlnAspAspTyrAlaAlaLysAspLeuAsnGlyLeu 291
Db      676  TTGGAGTTCCTCCGATTACCT-----CAGAAAAGACTGCCAGAACGCAAGCTTGA 729
Qy      292  SerGlyValArgAsnIleLeuTyrPseValGlyGly--GlyGlyAlaSerAlaLeuVal 310
Db      730  GAATTATTCGAAATATCTCTAGAAAGATCGTGGCGGATTCGACAGAAACAATGATGAA 789
Qy      311  GlyGlyLeuThrArgGlu-----LysPheMetAsnGlnAlaPheLeuAspMetThr 327
Db      790  TCGGGTTTAAACATGATGACTTTCGCGATTTCAGACGATTCAGCAACCACTTACAA 849
Qy      328  AlaAlaGlnGlnValAspLeuPheValAlaThrPseAsnIleProAlaGluSerPhe 347
Db      850  GTTCGACAGACGCAACGATTTTGAAGCTCAAGCAACGCT-----TCTTCGCTGT 903
Qy      348  GluValTyrGluValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLysLysProHis 367
Db      904  GCCACTTACTTAGCTGTATGCTTATGCTGATAGCGGAGATTCGCTCAACGCAACCTGCT 963
Qy      368  LeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnGlnAlaLysValMetAlaMet 387
Db      964  TTTAATTCGTCAAGCAGAACAAATGCTGCTGCTGCGCAAGCGCCAA----- 1011
Qy      388  GluIleProAlaMetLeuTyrAspThrArgAsnAsnTyrGluIleAspPheGlyLeuGlu 407
Db      1012 -----GATTCATCTTTCGAA 1026
Qy      408  ArgGlyLeuGlnCysAlaLeuLeuIleGlyLysValAspGluCysArgMetTyrLeuGlyLeu 427
Db      1027  CAGTCGTATGCTGCTTACTTTTGGGCAAACTGAGAGCAACCTCGTGTTCGAACTT 1086
Qy      428  AspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPheValLeuGluAsnSer 447
Db      1087  AGTCAGAG-----TAGGAA-----GCTTATGCTTTCGCGGAAAAATCT 1128
Qy      448  AsnArgAspAspAsnAspAspLeuProGlyLeuCysLysLeuLeuGluThrTyrLeuAla 467
Db      1129 -----CAGACTCTCCAGATTGTTTACCGGCTCTGTCTTATTCAGAGAACGTCGCTCA 1185
Qy      468  GlyValValPheProArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr 487
Db      1186  CACGAACTCTTCCCAATTCGAGATTTAGCAAAACAGCAAGCAAGCTTCCTTAAAGATTAC 1245
Qy      488  TyrAspAspProMetValLeuSerTyrLeuGlnArgValGluValValGlnGlySerPro 507
Db      1246  TTGCTTACCAACACAGGTGCAAGCTTATTTCGAA----- 1278
Qy      508  LeuAlaAlaAlaIleThrMetAlaArgIleGlyValGluHisValLysAlaSerAlaMet 527
Db      1279 -----GCACTGCCCACTGATGCCCAACACTATGATGAGGCTGTAATTAACCCCAAG 1332
Qy      528  GlnAlaLeuGlnLysValPheProSerArgTyrThrAspArgAsnSerAlaGluProLys 547
Db      1333  TATTTTCCCAAGCGCAAGCAAGAAATATCATTTTCATACCAATTCACATTAACCTTCA 1392
Qy      548 -----AspValGlnGluThrValPheSerVal 556
Db      1393  GCGTCATTATATCAACAGACAGTACTAAACCAAGATTCGCAAGAAACA----- 1440
Qy      557  AspProValGlyAsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAla 576
Db      1440 ----- 1440
Qy      577  ValArgProSerGluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGlu 596

```

```

Db      1441 -----CAAACAAA-----GAAACCTTGAATTCACAACTTTCACACCTATGTG 1488
Qy      597  SerSerValAspGluThrThrValGluMetSerValAlaAspMetLeuLys----- 613
Db      1489  AGTTCAATCGAGATTAATAATCAAGAGTTTCCTGCTGCTGAAAGATGACAGAGTACT 1548
Qy      614 -----GluAlaSerValLysIleLeuAlaAlaGly----- 623
Db      1549  AATCAGCATTTGAACGGTTCAGCTAAGAGTGTGCATCTGTGTCATTAACAAAACGTAAG 1608
Qy      623 ----- 623
Db      1609  CGAGAAAACCTTACTCTCATCTGCTAGCCGAGCGGTATACCAATATGCTCCTCATTTCT 1668
Qy      623 ----- 623
Db      1669  CGTGTCCCGAGAGCGGCGAATTTTGGCAACACATAGAAAGTAAACACGCGTGTGA 1728
Qy      624 -----ValAlaIleGlyLeuIleSerLeu----- 631
Db      1729  TGGAGAGTGTATTTCTTGTGTGACATATTAATGTTTGGGTATTAAGCCACAACT 1788
Qy      632 -----PheSerGlnLysTyrPheLeuLysSerSerSerPheGlnArgLysAspMet 649
Db      1789  TTTGATGTTTAAATAATCTGTTTCTTCTCAACCTTTCGCTGATCTACAGTGTGT 1848
Qy      650  ValSerSerMetGluSerAspValAlaThrIleGlySerValArgAlaAspAspSerGlu 669
Db      1849  GTACAAATAAACACACACCGCTTACCATTCGCGATCCAAATGAAACCAAGATTCGAA 1908
Qy      670  AlaLeuProArgMetAspAlaArgThrAlaGluAsnIleValSerLysTyrGlnLysIle 689
Db      1909  GAAGCCCTTTTAACAATGACAGAG--GCAGAAAGATTAATTCACACTTGTTATCTAAC 1965
Qy      690  LysSerLeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGly 709
Db      1966  AAAGCCGACGCTTTAGGCGCCCAATCAAGATTAATTAATTAAGCAAAATTTTAACTGCT 2025
Qy      710  ArgMetLeuLysIleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyr 729
Db      2026  TCAGCTTATCTCAATGCGCA-----CTGATTCCTCAACAGATTAAGTTAGCAATCGCTAC 2082
Qy      730 -----AspTyrThrLeuLeuLysLeuSerValAspSerValThrValSerAlaAsp 746
Db      2083  CGCAAGTTCACACATAGTTTGAAGATAGAAATCTGTTGAGAAATGCTTTATTTGACAGAT 2142
Qy      747  GlyThrArgAlaLeuValGlnAlaThrLeuGlnGluSerAlaCysLeuSerAspLeuVal 766
Db      2143 -----CTGCCGACGATGAGAGCTACGCTAAAGAAAGTACGCACTTATGAAATATAT 2196
Qy      767  HisProGluAsnAsnAlaThrAspValArgThrTyrThrThrArgTyrGlyValPheTyr 786
Db      2197  CAGTTTAAATACTTCTTAAGAT-----AAATTAAGATTCGGTATGACTTGAATTGCA 2250
Qy      787  SerLysSerGlyTyrLysIleThrGlnGlySerValLeu 799
Db      2251  GAACGAGTAAATGCGCTATTCAGAGTACATCTGTGTA 2289

```

Search completed: February 21, 2006, 13:05:15
 Job time : 1228 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 21, 2006, 13:17:07 ; Search time 1565 Seconds
(without alignments)
4232.442 Million cell updates/sec

Title: US-10-600-070B-2

Perfect score: 4063
Sequence: 1 MEALSHVIGLSPFQCRRLP.....YEVFMKSGMKITGSVLA 801

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWB.spool/US10600070/runat_21022006_112404_8748/app_query.fasta_1
-DB=Published Applications NA Main -OPMT=fastcap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1
-MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ads805h
-USER=US10600070 @CGN 1.1 1026 @runat_21022006_112404_8748 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US0A_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4063	100.0	2406	7 US-10-600-070-1	Sequence 1, Appl
2	4051	99.7	2406	7 US-10-600-070-128	Sequence 128, App
3	4051	99.7	2637	7 US-10-600-070-130	Sequence 130, App
4	4051	99.7	2679	8 US-10-739-930-227	Sequence 227, App
5	4049	99.7	2406	7 US-10-600-070-9	Sequence 9, Appl
6	3837.5	94.4	3667	7 US-10-600-070-3	Sequence 3, Appl
7	3823.5	94.1	3667	7 US-10-600-070-10	Sequence 10, Appl

8	1775.5	43.7	2283	7 US-10-600-070-126	Sequence 126, App
9	1134	27.9	1146	7 US-10-424-599-129007	Sequence 129007, A
10	973	23.9	1411	8 US-10-425-115-81853	Sequence 81853, A
11	960	23.6	1703	7 US-10-437-963-69933	Sequence 69933, A
12	856	21.1	1536	8 US-10-425-115-57452	Sequence 57452, A
13	784	19.3	561	7 US-10-600-070-132	Sequence 132, App
14	733.5	18.1	660	7 US-10-600-070-184	Sequence 184, App
15	718	17.7	660	7 US-10-600-070-135	Sequence 135, App
16	546	13.4	622	7 US-10-600-070-146	Sequence 146, App
17	541	13.3	537	7 US-10-600-070-143	Sequence 143, App
18	529	13.0	607	7 US-10-021-323-3562	Sequence 3562, App
19	524.5	12.9	2307	7 US-10-600-070-166	Sequence 166, App
20	503	12.4	608	7 US-10-600-070-137	Sequence 137, App
21	501	12.3	552	7 US-10-021-323-3636	Sequence 3636, App
22	499	12.3	1039	7 US-10-424-599-35059	Sequence 35059, A
23	498.5	12.3	2370	7 US-10-600-070-193	Sequence 193, App
24	497	12.2	2400	7 US-10-600-070-163	Sequence 163, App
25	482	11.9	2130	7 US-10-437-963-69932	Sequence 69932, A
26	480	11.8	491	7 US-10-600-070-174	Sequence 174, App
27	474.5	11.7	3554	7 US-10-437-963-53316	Sequence 53316, A
28	472	11.6	446	7 US-10-600-070-151	Sequence 151, App
29	465	11.4	545	7 US-10-600-070-175	Sequence 175, App
30	453.5	11.2	647	7 US-10-600-070-185	Sequence 185, App
31	448	11.0	871	7 US-10-600-070-153	Sequence 153, App
32	446	11.0	2469	7 US-10-600-070-4	Sequence 4, Appl
33	446	11.0	2469	7 US-10-600-070-161	Sequence 161, App
34	443	10.9	652	7 US-10-600-070-186	Sequence 186, App
35	435.5	10.7	647	7 US-10-767-701-4069	Sequence 4069, App
36	431	10.6	2857	7 US-10-600-070-172	Sequence 172, App
37	429	10.6	2145	7 US-10-600-070-168	Sequence 168, App
38	425	10.5	2022	7 US-10-600-070-191	Sequence 191, App
39	413	10.2	2511	7 US-10-600-070-189	Sequence 189, App
40	409.5	10.1	1032	7 US-10-767-701-9474	Sequence 9474, App
41	403	9.9	2151	7 US-10-600-070-159	Sequence 159, App
42	391.5	9.6	479	7 US-10-600-070-150	Sequence 150, App
43	390	9.6	460	7 US-10-600-070-187	Sequence 187, App
44	387.5	9.5	439	7 US-09-732-627A-2154	Sequence 2154, App
45	381.5	9.4	420	7 US-10-600-070-176	Sequence 176, App

ALIGNMENTS

RESULT 1
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseeryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Kokharova, Olga A.
; TITLE OF INVENTION: Plaacid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Alignment Scores:
Pred. No.: 0
Score: 4063.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 7
Gaps: 0
Length: 2406
Matches: 801
Conservative: 0
Mismatch: 0
Indels: 0

US-10-600-070B-2 (1-801) x US-10-600-070-1 (1-2406)

Qy 1 MercIuAlaIeuSerHisValIGIYIleGIYleuSerProphelIleuCyArgLeuPro 20
Db 1 ATGGAAGCTCTGAGTCACGTGGCATGTGTCTCCCATTCATTAATTCGCAATTACCA 60
Qy 21 ProAlaThrThrIyIleuValArgArgSerHisAAsnThrSerThrIleCySerAlaSer 40
Db 61 CCGGCGACGCAAAAGCTCCGACGTAGCCACAACCTCTAACACTATCTGCTCGGCAGC 120
Qy 41 LysThrAlaAAspArgLeuLeuSerAAspPheAAspPheThrSerAAspSerSerSer 60
Db 121 AAATGGGCCGACCGCTTCTCTCCGACTCAATTTCACTCCGATTCCTCTCTCTCC 180
Qy 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerIeuProProserIleAAspArgPro 80
Db 181 TTGGCCACCGCCACACCCACCGCACCTCTGTCCTCTGCGACCATATTAATGACGCC 240
Qy 81 GluArgHisValProIleProIleAAspPheTyArgIleValIleuGIYAlaGlnThrHisPhe 100
Db 241 GAACGCCACGTCGCCATCCCATTTGATTTCTACCAAGTATTAAGAGCTCAAAACATTC 300
Qy 101 LeuThrAAspGIYIleArgArgAlaPheGIYAlaArgValSerIySProProGlnPheGIY 120
Db 301 TTAACCGATGGAATCAGAAAGACATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGT 360
Qy 121 PheSerAAspAAspAlaIeuIleSerArgArgGlnIleIleuGlnAlaAlaCySGIuThrLeu 140
Db 361 TTCAGCCACACGCTTTAAATCAGCGGAGACAGATCTTCAAGCTGCTTGCGAAACTCTG 420
Qy 141 SerAAspProArgSerArgArgGIYTyArgGIYIleuGIYAlaIleuAAspArgIleuAlaThr 160
Db 421 TCTAATCTCGCTCTTAAGAAAGAGTACATGAAGTCTTCTGTGATGATGAAGAGCTACA 480
Qy 161 ValIleThrAAspValProThrAAspIyValProGIYAlaIleuCySValIleuGlnIleGIY 180
Db 481 GTCATCACTGATGTCTCTGGGATAGATGCTCTGCGGCTCTCTGTGATGTCAGAAAGGT 540
Qy 181 GIYGIuThrArgIuIleValIleuArgValGIYGIYAlaIleuIleuIySGIuArgLeuProIyS 200
Db 541 GGTAGAGCTAGAGATGATCTTGGGGTGTGAGGCTCTGCTTAAGAGAGGTTCCCTAAG 600
Qy 201 SerPheIySGIuAAspValIleuValIleuAlaIleuAlaPheLeuAAspValSerArgAAsp 220
Db 601 TCGTTTAAGCAAGATGTGTGTTTATGATAGCGCTTGCTTGCTCGATGTCTGAGGGAT 660
Qy 221 AlaMetAlaIleuAAspProProAAspPheIleThrGIYTyArgIuPheValGIYGIYAlaIleu 240
Db 661 GCATGTGCAATGTGATCCACTGATTTTATTACTGTTATGAGTTGTGTGAGGAAGCTTTG 720
Qy 241 LysIleuIleuGlnIleuGIYAlaIleuSerSerIleuAlaProAAspIleuArgAlaGlnIleAAsp 260
Db 721 AAGCTTTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Qy 261 GIuThrIleuGIYIleuIleuThrProArgTyArgValIleuGIYAlaIleuProIleuGIY 280
Db 781 GAGCTTTTGAAGAGATCACTCCGCTTATCTTGTGAGCTACTTGGTTTACCGCTTGCT 840
Qy 281 AAspAAspTyAlaAlaIyArgLeuAAspGIYLeuSerGIYValArgAAsnIleuIleuTrpSer 300
Db 841 GATGATTAACGCTCGGAAAAGACATAATGTTTAAAGCGGTGCGGAAATATTTGTGGCT 900
Qy 301 ValGIYGIYGIYAlaSerAlaIleuValGIYGIYLeuThrArgGIYIySPhMetAAsn 320
Db 901 GTTGAGAGAGGTGAGCATCAAGCTCTGTGTGGGGTGTGACCGCGAGAAATTTATGAT 960
Qy 321 GIYAlaPheLeuAAspMetThrAlaAlaGIYIleuValAAspLeuPheValAlaThrProSer 340
Db 961 GAGCGCTTTTACCAATGACAGCTGCTGACAGAGTGAATCTTTTGTAGTACACCGCAAGC 1020
Qy 341 AAsnIleProAlaIleuSerPheGIYValTyArgIuValAlaIleuAlaIleuValAlaGlnAla 360
Db 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGCTGTGTGCTCAAGCT 1080

Qy 361 PheIleGIYIySArgProHisIleuIleuGlnAAspIySGIuPheGlnIleuGln 380
Db 1081 TTTATGTGTAAGAGCACCTTTTACAGAGATGTGATTAAGCAATTCACAGCAATTCAG 1140
Qy 381 GlnAlaIySValMetAlaMetGIYIleProAlaMetLeuTyArgPheArgAAsnAAspTrp 400
Db 1141 CAGGCTAAGGTAAATGGCTATGAGATTCCTGCATATGTATGATACAGCAATTAATGG 1200
Qy 401 GIYIleAAspPheGIYIleuGIYAlaArgGIYLeuCyAlaIleuIleuIleGIYIySValAAspGIY 420
Db 1201 GAGATGACTTCGCTCTTAAGAAAGGAGACTGTGCACTGCTTATAGGCAAAAGTTGATGA 1260
Qy 421 CyArgMetTrpIleuGIYLeuAAspSerGIYAAspSerGlnTyArgAAspProAlaIleVal 440
Db 1261 TGCCTATGTGTGGCTTAGACAGTGAAGATTCACAATATATGAAATCCAGCTAATGTG 1320
Qy 441 GIYIlePheValIleuGIYIleuSerAAsnAlaArgAAspAAspAAspIleuProGIYLeuCyIyS 460
Db 1321 GAGTTGTGTTTGAAGAAATTCMAATCGATGACAAATGATATCTCCCTGAGCTATGCAAA 1380
Qy 461 LeuIleuGIYThrTrpIleuAlaGIYValIlePheProArgPheArgAAspThrIySArgIyS 480
Db 1381 TTGTTGAAACCTGCTGTGAGGAGGCTGTCTTCTTAAGTTCAAGACACCAAAAGTAA 1440
Qy 481 LysPheIySLeuGIYAAspTyTyArgAAspPrometValIleuSerTyIleuGIYArgVal 500
Db 1441 AAATTTAACTCGGGGACTATGATGATCTATGATGATCTTATGATGATCTTGAAGAGCTG 1500
Qy 501 GIYValIleGIYIleIySProIleuAlaAlaAlaIleuThrMetAlaArgIleGIYAlaGIY 520
Db 1501 GAGTGTGTTCAAGGCTTCTCTTAACTGCTGCTGCAACTATGCGAAAGATTTGAGCCGAG 1560
Qy 521 HisValIySAlaSerAlaMetGlnAlaIleuGlnIySValPheProSerArgTyThrAAsp 540
Db 1561 CATGTAAAGCTAGCTGTATGACAGCACTGCAGAAATTTTCTTCCGCTATACAGAT 1620
Qy 541 ArgAAspSerAlaIleuProIySArgValGIYIleuThrValPheSerValAAspProValGIY 560
Db 1621 AGAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGATGATCTCTGTGGT 1680
Qy 561 AAspAAsnValGIYArgAAspGIYIleuProGIYValPheIleAlaGIYAlaValAAspProSer 580
Db 1681 AACATGTAGCCGTGATGTGAGCTGTGTCTTATGCAAGAGCTGTAAGACCTCT 1740
Qy 581 GIYIleAAspPheGIYThrAAspAAspTyAlaIleArgAlaGIYAlaIleuGIYIleuSerIyS 600
Db 1741 GAATACTTTGAACCTAATGATTTATGCAATTCAGAGCTGCGGCTCAAGAGATAGCTTGAT 1800
Qy 601 GIYThrThrValGIYIleuSerValAlaAAspMetIleuGIYAlaIleuValIyS 620
Db 1801 GAATCTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Qy 621 AlaAlaGIYAlaAlaIleGIYIleuIleSerLeuPheSerGlnIySArgPheIleuIyS 640
Db 1861 GCTGCTGCTGTGCAATTTGACATGATTTCACTGTTCAGCCAGAAATTTTCTTAAAGC 1920
Qy 641 SerSerSerPheGlnArgIySArgMetValSerSerMetGluSerAAspValAlaThrIle 660
Db 1921 AGCTCATCTTTTCAACCGCAAGATATGTTTCTTCTATGAAATCTGATGTCTGATACATA 1980
Qy 661 GIYSerValArgAlaAAspAAspSerGIYAlaIleuProArgMetAAspAlaArgThrAlaGIY 680
Db 1981 GGGCTAGCTAGAGTGAAGTTCAGAGGACCTTCCAGAAATGATGTCTAGAGCTGACAG 2040
Qy 681 AAsnIleValSerIySArgIySAlaIleuSerIleuAlaPheGIYProAAspHisArgIle 700
Db 2041 AATATGATATCCAGTGTGAGAAAGATTAAGTCTCTGCTTTTGGGCTGTATCACCGCAT 2100
Qy 701 GIYIleuIleuProGIYValIleuAAspGIYArgMetLeuIySArgPThrAAspArgAlaAla 720
Db 2101 GAAATGTACCAAGGTTTGTGATGAGGCAATGTGCAAGATTTGACATGACAGAGCAAGCT 2160
Qy 721 GIYThrAlaIleuGIYIleuValTyArgAAspTyThrIleuIleuIySLeuSerValAAspSer 740


```

Db      2161  GAAACTCGCAGCTGGGTGGTTATGATATACAGTGTGAACATATCTGTGACAGT 2220
      741  ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluSerAla 760
      2221  GTGACATCTTCAGACATGGAACCGGTCTGTGTGAAAGCACTCGAGAGACTGTGCT 2280
      761  CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
      2281  TGTCTATCTGATTTGGTTTCATCCGAAACATCTCTGATGTCAAGACTTACACACA 2340
      781  ArgTyrGluValPheTyrSerLysSerGlyTyrLysIleThrGluGlySerValLeuAla 800
      2341  AGATACGAAGTTTCTGTGTCAGACTCAGGGTGAAATCACTGAAGGCTGTGTTCTTGA 2400

      801  Ser 801
      2401  TCA 2403

RESULT 2
US-10-600-070-128
; Sequence 128, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitcha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

Alignment Scores:
Pred. No.: 0          Length: 2406
Score: 4051.00       Matches: 799
Percent Similarity: 99.8%  Conservative: 0
Best Local Similarity: 99.8%  Mismatches: 2
Query Match: 99.7%  Indels: 0
DB: 7          Gaps: 0

US-10-600-070b-2 (1-801) x US-10-600-070-128 (1-2406)
      1  MercGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGluLeuCyArgLeuPro 20
      1  ATGGAAGCTCTGAGTCACTCGGCACTTGTCTCTCCCATTCACATTAATCCGATTACCA 60
      21  ProAlaThrThrIleValLeuArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
      61  CCGGCACGACAAAGCTCCGACGAGCAACAACCTCTCAACTACTCTCTCCGCGCAGC 120
      41  LysTyrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
      121  AATATGGCCGACGCTCTCTCTCTCGACTTCATTTTCCATTCCTCCGATTCCTCTCTCC 180
      61  PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
      181  TTTCGCACACCGCCACACACGCGCACTCTGTCTCTCCGACCATCTATTGATCCGCC 240
      81  GluArgHisValProIleProIleAspPheTyrGluValLeuGlyIleGlnThrHisPhe 100
      241  GAAAGCGACGTCCTCCCATCCCATTTGATTCTTACCAAGTATTAGAGCTCAAAACATTC 300
      101  LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProGlnPheGly 120

```

```

      301  TTAACCGATGAAATCAGAGAGCATTCGAGCTAGGCTTTCGAAACCGCCGCAATTCGCT 360
      121  PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGlyThrLeu 140
      361  TTCAGCAGCAGCGTTTAAATCAGCCGAGACAGATTCTTCAAGTGGCTTGGAACCTTG 420
      141  SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
      421  TCTAATCTCGCTCTAAGAGAGTACATGAAGAGCTTCTTGAATGATGAAGAGCTACA 480
      161  ValIleThrAspValProTyrAspLysValProGlyAlaLeuCysValLeuGlnGly 180
      481  GTCATCACTAGATGTTCTTGAGATAGAGTTCCTGCTCTCTGTGATTCAGAAAGGT 540
      181  GlyGluThrGluIleValLeuArgValGlyGluAlaLeuLeuLysGluValArgLeuPro 200
      541  GGTGAGACTAGATGATGTTCTTGGGTGGTGGAGCTTGTCTTAAAGAGAGGTGCTTACG 600
      201  SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
      601  TCGTTTAAAGCAAGATGCTGTTTATGATGCGCTGCTTCCGATGCTCGAGGGAT 660
      221  AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
      661  GCTATGCGCATTTGATCACTGATTTTAACTGTTATAGATTGATGATTTGTTGAGAAAGCTT 720
      241  LysLeuLeuGlnGlnGluGluGlyAlaSerSerLeuAlaProAspLeuAspAlaGlnIleAsp 260
      721  AACCTTTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
      261  GluThrLeuGlnGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGly 280
      781  GAGACTTTGAGAGAGATCACTCCGCTTATGCTTGAGAGCTTACCTGCTTACCGCTTGT 840
      281  AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValAlaArgAsnIleLeuTyrSer 300
      841  GATGATTACGCTCGAAGAAAGACTAAATGTTTAAAGCGGTGCGGAATATTGTGTGCT 900
      301  ValGlyGlyGlyValAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsn 320
      901  GTTGAGAGAGGTGAGATCAGCTCTGTTGGGGGTGTGACCGGTGAGAGATTATATAT 960
      321  GluAlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrProSer 340
      961  GAGCGCTTTTACGATGAGACGCTGCTGAGCAGAGTTCATTTTGTAGCTTACCCAGC 1020
      341  AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAla 360
      1021  AATATTCAGCAGACAGTCAATTGAAGTTTACGAAGTTGCACTGTCTTGTGGCTCAAGCT 1080
      361  PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
      1081  TTTATTTGTAAGAGCAGCACCTTTTACAGAGTCTGATTAAGATTCACGAACTTCAG 1140
      381  GluAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTyr 400
      1141  CAGGCTTAAGTAAATGCTATGAGATTCCTGCGATGTTGATATACACGAAATATATGG 1200
      401  GluIleAspPheGlyLeuGlnArgGlyLeuCyAlaLeuLeuIleGlyLysValAspGlu 420
      1201  GAGATTAAGCTTGGTCTTAAGAAAGGGAGCTGTGCACTGTTATAGGCAAAAGTTGATGA 1260
      421  CysArgMetTyrLeuGlnLysLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
      1261  TGCCTGATGTGTTGGCTTAAACAGTGAAGATTCAAAATATGGAATTCACGATATTTG 1320
      441  GluPheValLeuGlnLysSerAsnArgAspAspAspAspLeuProGlyLeuCysLys 460
      1321  GAGTTGTTTGGAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGACATATGCAAA 1380
      461  LeuLeuGluThrTyrLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
      1381  TTGTTGAAACCTGTGTGGCAGGGTGTCTTCTTCTAGTTTCAGAGACACCAAAAGATAAA 1440

```

```
QY 481 LysPheIyLeuGlyAspTyrTrpAspPromeValLeuSerTyrLeuGlyArgVal 500
DB 1441 AATTTAACTCGGGGACTATGATGATCTTATGATTGATTGAGTAAAGGAGTG 1500
QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGlu 520
DB 1501 GAGGTAGTTCAGGGTCTCTTATAGCTGCTGCGACCTATGGCAAGATTGGAGCCGAG 1560
QY 521 HisValIyValAspSerAlaMetGlnAlaLeuGlnIyValPheProSerArgTyrThrAsp 540
DB 1561 CATGTGAAGCTAGTGTCTATGACAGGCACTGCAGAAAGTTTCTTCCCTATACAGAT 1620
QY 541 ArgAsnSerAlaGluProIyAspValGlnGluThrValPheSerValAspProValGly 560
DB 1621 AGAACTCGGCTGAAACCAAGAGTGTGCAAGACAGATGTTTATGTTAGTATCCGTGGT 1680
QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
DB 1681 AACAAATGTAGGCGGTGATGTGAGCTGTGTCTTATTTGCAAGAGCTGTAAAGCCCTCT 1740
QY 581 GluAsnPheGlnThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerValAsp 600
DB 1741 GAAACTTGTAAACTATATGATTATGCAATTCAGCTGGGCTCTCAGAGATAGCGTTGAT 1800
QY 601 GluThrThrValGluMetSerValAlaAspMetLeuIySerGluValSerValIyLeu 620
DB 1801 GAAACTACTGTGTGAAAGTCCGTGTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
QY 621 AlaAlaGlyValAlaIleGlyIyLeuIleSerLeuPheSerGlnIyTyrPheLeuIySer 640
DB 1861 GCTGCTGTGTGGCAATTGACATGATTTCACTGTTCACCGCAAGATTTTCTTAAAGC 1920
QY 641 SerSerSerPheGlnArgIyAspMetValSerSerMetGluSerAspValAlaThrIle 660
DB 1921 AGCTCATCTTTTCAACCGCAAGATATGTTTCTTCTATGCAATCTGATGTGCTACCATTA 1980
QY 661 GlySerValArgAlaAspAspSerGlnAlaLeuProArgMetAspAlaArgThrAlaGlu 680
DB 1981 GGGTCATCTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAAGATGATGCTAGAGCTGCAGAG 2040
QY 681 AsnIleValSerIySerIyProGlnIyIleIySerSerLeuAlaPheGlyProAspHisArgIle 700
DB 2041 AATTTAGTATTCAGAGTGCAGAGATTTAAGTCTGTGCTTTTGGCGCTGATCAACGCATA 2100
QY 701 GluMetLeuProGlnValIleuAspGlyArgMetLeuIyIleTyrThrAspArgAlaAla 720
DB 2101 GAATGTTCACAGAGTTTGGATGGGCGAATGCTGAAGATTGTGACTGCAGAGCAGCT 2160
QY 721 GluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuIySerValAspSer 740
DB 2161 GAAACTCGCAGCTTGGGTGGTTATGATTATACACTGTGAAACTATCTGTGACAGT 2220
QY 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 760
DB 2221 GTGACAGTCTCAGAGATGAAACCCGTCTCTGTGTGAAGCACTCTGGAGAGATCTGCT 2280
QY 761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
DB 2281 TGCTATCTGATTTGGTTCACTCAGAAACAACTGCTACTGATGTCAAGAACTACACAA 2340
QY 781 ArgTyrGlnValPheTyrPheSerIySerGlyTyrIyValIleThrGluIySerValLeuAla 800
DB 2341 AGATACCAAGTTTCTGTCTCAAGTCAAGGTCGAAATCACTGAAGGCTGTGTTCTTGCA 2400
QY 801 Ser 801
DB 2401 TCA 2403
```

RESULT 3
US-10-600-070-130
; Sequence 130, Application US/10600070
; Publication No. US20040139500A1

```
; GENERAL INFORMATION:  
; APPLICANT: Osteoryoung, Katherine W.  
; APPLICANT: Vitha, Stanislaw  
; APPLICANT: Koksharova, Olga A.  
; APPLICANT: Gao, Hongbo  
; TITLE OF INVENTION: Platelet Division and Related Genes and Proteins, and Methods of  
; FILE REFERENCE: MSU-08153  
; CURRENT APPLICATION NUMBER: US/10/600,070  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 130  
; LENGTH: 2637  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-600-070-130  
  
Alignment Scores:  
Pred. No.: 0 Length: 2637  
Score: 4051.00 Matches: 799  
Percent Similarity: 99.8% Conservative: 0  
Best Local Similarity: 99.8% Mismatches: 2  
Query Match: 99.7% Indels: 0  
DB: 7 Gaps: 0  
  
US-10-600-070B-2 (1-801) x US-10-600-070-130 (1-2637)  
  
QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyAspArgLeuPro 20  
DB 114 ATGAAGCTCTGAGTACGCGGCATGTCCTCCCAATTCAAATTATGCGAATTAACA 173  
  
QY 21 ProAlaThrThrIySerLeuArgArgSerHisAsnThrSerThrIleCysSerAlaSer 40  
DB 174 CCGCGACGACAAAGCTCCAGACTAGCCACACACACTTCAACTATCTCTCCGACG 233  
  
QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60  
DB 234 AATGGGCGGACCGTCTTCTCGAATTCATTCACCTCCGATTCCTCTCTCC 293  
  
QY 61 PheAlaThrAlaThrThrThrAlaThrIleValIleSerLeuProProSerIleAspArgPro 80  
DB 294 TTGCGCCACCGCCACACACACCGCAGCTGTCTCCGACACATCTATTTGATCGTCC 353  
  
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValIleGlnIyAlaGlnThrHisPhe 100  
DB 354 GAAGCCAGACGTCCTCCATCCCATTTGATTTTACACAGGATTTAGAGACTCAACACATTC 413  
  
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIyProProGlnPheGly 120  
DB 414 TTAAACGATGGAATCAGAAAGAGCATTCGAAGTATGAGGTTTCGAAACCGCCGAATTCGGT 473  
  
QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCysGluThrLeu 140  
DB 474 TTCAGGACACACGCTTATACGCGGAGACAAATCTTCAAGCTGCTTCGAAACTCTG 533  
  
QY 141 SerAsnProArgSerArgArgGluTyrAsnGlyIyLeuLeuAspAspGluGluAlaThr 160  
DB 534 TCTAATCTCGGTCTTAAGAGAGTACAAAGAGGCTTTCTTATGATGATGAAGAGTACA 593  
  
QY 161 ValIleThrAspValProTyrAspIyValProGlyAlaLeuCyValLeuGlnGly 180  
DB 594 GTATCTCATGATGTTCTTGGATTAAGTTCGTGTCTCTCTGTGATTTGAAGAGAGT 653  
  
QY 181 GlyIleThrGluIyIleValLeuArgValGlyGluAlaLeuLeuIySerGluArgLeuProIy 200  
DB 654 GGTGAGACTGAGATGTTCTTCCGGTTGTGAGGCTGTGCTTAAAGAGAGGTCCTTAAG 713  
  
QY 201 SerPheIySerGlnAspValValIleuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220  
DB 714 TCGTTTAAAGCAAGATGTGTTTATGATGAGCGCTTTCGATGTCTCAGATGCTCAGAGGAT 773  
  
QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
```

```

Db 774 GCTATGCAATTCAGTCCACCTGATTTATACCTGTTATGATTTGTTGAGGAAGCTTGG 833
Qy 241 LyeLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
Db 834 AACCTTTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 893
Qy 261 GluThrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 894 GAGACTTTGAGAGAGATCACTCCGCGTTAGTCTTGAGAGCTTGGCTTACCGCTTGGT 953
Qy 281 AspAspTyrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 300
Db 954 GATGATTAACCTGCGAAGAAAGCTTAATGTTTAAGCGGTGTGCGGAATATTTTGCTCT 1013
Qy 301 ValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 320
Db 1014 GTTGGAGAGAGTGAAGATCAGCTCTTGTGGGGGTTGACCCGTGAGAACTTATGAAAT 1073
Qy 321 GluAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
Db 1074 GAGCGCTTTTACGAATGACAGCTGCTGAGAGGTTGATCTTTTGTAGCTTACCCAGC 1133
Qy 341 AsnIleProAlaGlnSerPheGlnValTyrGlnValAlaLeuAlaLeuValAlaGlnAla 360
Db 1134 AATATTCCAGCAGAGTCAATTTGAAGTTTGAAGTTGCACTTGTCTTGTGGCTCAAGCT 1193
Qy 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnLeuGln 380
Db 1194 TTTATTGGTAAAGAACCACTTTTACAGAGTCTGATTAAGCAATTCAGCAACTTCAG 1253
Qy 381 GlnAlaLysValMetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAspAspTyr 400
Db 1254 CAGCGCTTAAGGTAATGGCTATGAGATTCCTGCGAGTGTATGATTAACGAAATATTTGG 1313
Qy 401 GluIleAspPheGlyLeuGlnArgGlyLeuGlnAlaLeuLeuIleGlyLysValAspGln 420
Db 1314 GAGATTGACTTGGTCTAGAAAGGAGACTGTGTGCACTGCTTATAGCAAAAGTTGATGAA 1373
Qy 421 CysArgMetTyrLeuGlnLysLeuAspSerGlnAspSerGlnTyrArgAspProAlaIleVal 440
Db 1374 TGCCGTATGTGTGGGCTTGAAGACGTGAGAGATTCACAAATATAGGAATCCAGCATTTGTG 1433
Qy 441 GluPheValLeuGlnLysSerAspArgAspAspAspAspAspAspAspAspAspAspAsp 460
Db 1434 GAGTTGTTTGAAGATTAATCAATCGATGACAAATGATATCTCCCTGAGCAATATGAAA 1493
Qy 461 LeuLeuGlnThrTyrPheLeuAlaGlyValAlaPheProArgPheArgAspThrLysAspLys 480
Db 1494 TTGTTGGAACCTGCTTGGGAGGGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1553
Qy 481 LysPheLysLeuGlnLysAspTyrTyrAspAspPheMetValLeuSerTyrLeuGlnArgVal 500
Db 1554 AATATTAACTCGGAGACTACTATGATGATCTTATGTTTGAAGTTTCTTCTTCTTCTTCTT 1613
Qy 501 GluValValGlnGlnLysSerProLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 520
Db 1614 GAGGTATGTCAGGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1673
Qy 521 HisValLysValAspSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
Db 1674 CATGTGAAGCTAGTCTATGACAGGCACTGCAAGAAATTTTCTTCTTCTTCTTCTTCTT 1733
Qy 541 ArgAspSerAlaGlnProLysAspValGlnGlnThrValPheSerValAspProValGly 560
Db 1734 AGAACTCTGGCTGACCAAGAGTGTGCAAGACAGTGTATGATGATGATGATGATGATGAT 1793
Qy 561 AsnAsnValGlyArgAspGlyGlnProGlyValPheIleAlaGlnAlaValAlaArgProSer 580
Db 1794 AACAAATGTAGGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
Qy 581 GluAsnPheGlnThrAsnAspTyrAlaIleArgAlaGlyValSerGlnSerValAsp 600

```

```

Db 1854 GAAACTTTGAACATAATGATTAAGCAATTCGAGCTGGGCTTCAGAGAGTGAAGCTTGTAT 1913
Qy 601 GluThrThrValGlnMetSerValAlaAspMetLeuLysGlnAlaSerValLysIleLeu 620
Db 1914 GAAACTTACTGTTGAAAGTCCGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1973
Qy 621 AlaAlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnTyrTyrPheLeuLysSer 640
Db 1974 GCTGCTGTGTGGCAATTTGACATTTTCACTGTTCCAGCCAGAAATTTTCTTAAAGC 2033
Qy 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAspValAlaThrIle 660
Db 2034 AGCTTACTTTTCAACGCAAGATATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2093
Qy 661 GlySerValAlaArgAlaAspAspSerGlnAlaLeuProArgMetLeuAlaArgThrAlaGln 680
Db 2094 GGGTCAGTCAAGACTGACATTCAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGAT 2153
Qy 681 AsnIleValSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
Db 2154 AATATTGATTCAGAGTGGAGAAAGATTAATGCTCTGCTTTTGGCTGATCACCGCAT 2213
Qy 701 GluMetLeuProGlnValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAla 720
Db 2214 GAATGTATACAGAGGTTTGGATGGCGAATGCTGAAGATTTGACAGAGCAGCT 2273
Qy 721 GluThrAlaGlnLeuGlnLysLeuValTyrAspTyrThrLeuLeuLysLeuSerValAspSer 740
Db 2274 GAAACTGCGCAGCTGTGGGTTGATTAATTAACCTGTGAAATCATCTGTTGACAGT 2333
Qy 741 ValThrValSerAlaAspGlyThrArgAlaLeuValGlnAlaThrLeuGlnGlnGlnGln 760
Db 2334 GTGACAGCTTCACAGATGAGAACCCGCTCTGTGTGAGAACCACTTGGAGAGTCTGCT 2393
Qy 761 CysLeuSerAspLeuValHisProGlnAsnAsnAlaAlaThrAspValAlaArgThrTyrThr 780
Db 2394 TGCTATTCGATTTGGTGTATCCAGAAACCAATGATGATGATGATGATGATGATGATGATGAT 2453
Qy 781 ArgTyrGlnValPheTyrSerLysSerGlyTyrLysIleThrGlnGlnSerValLeuAla 800
Db 2454 AGATTAAGAAATTTTGTGCTCAAGTCAAGGTGAGAAATCACTGAGGCTCTGTTTGTGA 2513
Qy 801 Ser 801
Db 2514 TCA 2516

RESULT 4
US-10-739-930-227
; Sequence 227, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 227
; LENGTH: 2679
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23ABP03-CLUSTER13643_1
US-10-739-930-227

Alignment Scores:
Pred. No.: 0 Length: 2679
Score: 4051.00 Matches: 799
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 2
Query Match: 99.7% Indels: 0
DB: 8 Gaps: 0

```

US-10-600-070b-2 (1-801) x US-10-739-930-227 (1-2679)

QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuGlyArgLeuPro 20
 Db 115 ATGGAAAGCTCTGAGTACGTCGGCATTGGTCTCTCCCATTTCCAAATTATGCCCATTACCA 174
 QY 21 ProAlaThrThrIysLeuArgArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
 Db 175 CCGGCACAGCAAGAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCTCCGCCAGC 234
 QY 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
 Db 235 AATATGGCCGACCGCTCTCTCCGACTTCATTTGACCTCCGATTCTCTCTCTCCCTCC 294
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 Db 295 TTCGCCACCGCCACACACCGCCACTCTGCTCTCCGCCACCATCTATTTGATCGTCCC 354
 QY 81 GluArgHisValProIleProIleAspPheTyrGluValLeuGluValAsnThrHisPhe 100
 Db 355 GAACGCAAGTCCCATCCCATTTGATTCTTACCAAGTATTAGAGCTCAAAACAATTTC 414
 QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
 Db 415 TTATCCGATGAGATCAGAAAGCATTGGAAGCTTAGGGTTTGAACCGCGCAATTCGGT 474
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysGluThrLeu 140
 Db 475 TTCAGGACGACCCCTTTAATCAGCCGAGACAGATTCTTCAGCTGCTCGCAAACTCTG 534
 QY 141 SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
 Db 535 TCTAATCTCTGGCTAGAGAGAGTACAAAGAGCTTCTTGATGATGAAGAAGCTACA 594
 QY 161 ValIleThrAspValProTrpAspLysValProGlyAlaLeuGlyValLeuGlnGly 180
 Db 595 GTCATACGTAGTCTCTGGATTAAGTTCCTGGTCTCTCTGTGATTTGCAAGAAAGGT 654
 QY 181 GlyGluThrGluIleValIleLeuArgValGlyGluAlaLeuLeuLysGluValLeuProLys 200
 Db 655 GGTGAGACTGAGTAGTCTTCCGGGTGGAGGCTCTGCTTAAGAGAGGTGGCTTACG 714
 QY 201 SerPheLysGlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 Db 715 TCGTTAAGCAAGATGGTGTATTAGTTATGAGCGCTTCCGTTCTCAATGCTCGAGGAT 774
 QY 221 AlaMetAlaLeuAspProProAspPheIleThrGlyTyrGluPheValGluGluAlaLeu 240
 Db 775 GCTATGGCATTTGGATCCACTGATTTTATTAACGTGTTATGAGTTGTTGAGAAAGCTTTG 834
 QY 241 LysLeuLeuGlnGluGluGlyAlaSerSerLeuAlaProAspLeuAlaGlnIleAsp 260
 Db 835 AACCTTTTACAGAGAGAGAGAGCAAGTACCTTGGCACCGGATTTTACCTGCACAAATTGAT 894
 QY 261 GluThrLeuGluGluIleThrProArgTyrValLeuGluLeuLeuGluGlyLeuProLeuGly 280
 Db 895 GAACCTTTGAAAGATCACTCCGCGTTATGTCTTGAGACTACTGGCTTACCGCTTGGT 954
 QY 281 AspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSer 300
 Db 955 GATGATTACGCTCGAAGAAAGCTTAATGGTTTAAAGCGGTGCGGAATATTTGTGCTCT 1014
 QY 301 ValGlyGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlyLysPheMetAsn 320
 Db 1015 GTTGGAGAGGTGAGCATCAGCTCTTGGGGGTGTGACCCCTGGAAGTATTATTAAT 1074
 QY 321 GluAlaPheLeuArgMetThrAlaAlaGluGluValAspLeuPheValAlaThrProSer 340
 Db 1075 GAGCGCTTTTACCAATGACGCTGCTGAGGAGGTATCTTTTGAAGCTTACCCCAAGC 1134
 QY 341 AsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuValAlaGlnAla 360

Db 1135 AATATTCACGACAGTATTGAAGTTTACGAAGTTGCACTTGCTCTTGCTCAAGCT 1194
 QY 361 PheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGln 380
 Db 1195 TTTATTTGGTAAAGCCACACTTTTACAGATGCTGATTAACAAATTTCCAGCAACTTACG 1254
 QY 381 GlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrp 400
 Db 1255 CAGGCTTAAGTATGCTATGAGATCTCTGCGATGTTGTATGATACAGGAATATATTGG 1314
 QY 401 GluIleAspPheGlyLeuGluValArgGlyLeuCysValAlaLeuLeuIleGlyLysValAspGlu 420
 Db 1315 GAGATAGACTTGCTGTATGAAGAGGACTGTGCACTGCTTATAGCCAAAGTTATGAA 1374
 QY 421 CysArgMetTrpLeuGlyLeuAspSerGluAspSerGlnTyrArgAsnProAlaIleVal 440
 Db 1375 TGCCGATGTGGTGGGCTTTACACATGAGATTCAATATAGAAATCCAGCTATTGGG 1434
 QY 441 GluPheValLeuGluLysAsnSerAsnArgAspAspAsnAspLeuProGlyLeuCysLys 460
 Db 1435 GAGTTTGTTTGGAGAAATCAATCGTATGACAAATGATGATCTCCCTGCACTATGCAAA 1494
 QY 461 LeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLys 480
 Db 1495 TTGTTGGAACCTGTGTGGCAGGGGTGTCTTCTTACGGTTACAGACACCAAGATTAA 1554
 QY 481 LysPheLysLeuGlyAspTyrTyrAspAspProMetValLeuSerTyrLeuGluArgVal 500
 Db 1555 AATTTTAACTCGGGACCTACATATGATTCATATGTTTGAAGTACTTGGAAAGAGTG 1614
 QY 501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgIleGlyValGlu 520
 Db 1615 GAGGTAGTTCAAGGTTCTCTTACGCTGCTGCGACTATGCAAGATTGAGCCGAG 1674
 QY 521 HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPheProSerArgTyrThrAsp 540
 Db 1675 CATGTAAAGCTAGTGTCTATGACGACATGCAAGAAATTTTCTCCGCTATACAGAT 1734
 QY 541 ArgAsnSerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
 Db 1735 AGAAATCGGCTGAACCCAGATGTCNAAGAGACAGTGTTTATAGTACATCCTGTTGGT 1794
 QY 561 AsnAsnValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSer 580
 Db 1795 AACCAATGATGACCGTATGATGATGAGCTGCTGCTTATTTATGCAAGACCTGTAAGACCCTCT 1854
 QY 581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
 Db 1855 GAAACCTTTGAACTATATATATTAATTCGAATTCAGCTGGGGGTCTCAGAGAGTACGTTGAT 1914
 QY 601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu 620
 Db 1915 GAACCTACTGTTGAATATGTCCTGCTGATATGTTAAAGAGCAAGTGTGAAGATCCTTA 1974
 QY 621 AlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSer 640
 Db 1975 GCTGCTGCTGGCAATTGACATTCATCTGTTCAAGCCAGCAAGATATTTCTTAAAGC 2034
 QY 641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
 Db 2035 AGCTTACTTTTAAAGCAAGATATGATGTTCTTCTATGATCTATGATCTGCTACATA 2094
 QY 661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
 Db 2095 GGGTCACTCAAGACTACATTCATTAAGACACTTCCAGAAATGATCTGATGCACTGCAGAG 2154
 QY 681 AsnIleValSerLysTrpGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIle 700
 Db 2155 AATATATGATCCAAAGTGGCAGAGATTAAGTCTCTGGCTTTTGGCGCTGATCAACCATTA 2214
 QY 701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleThrThrAspArgAlaAla 720
 Db 2215 GAAATGTTACCAAGAGTGTGTGATGGCGCAATGCTGAAGATTTGACCTGACAGACAGCT 2274

QY 721 G|uThrAlaGlnLeuG|yLeuValTyrAspTyrThrLeuLeuLyLeuSerValAspSer 740
 Db 2275 GAACTGGCGACGCTTGGGTTGGTTATATATACACTGTTGAACATCTCTGTTGACAGT 2334
 QY 741 ValThrValSerAlaAspG|yThrArgAlaLeuValGlnAlaThrLeuGlnG|uSerAla 760
 Db 2335 GTGACAGTCTCAGCAGATGAAACCGTGCTCTGGTGGAAAGCACTCGAGAGAGTGGCT 2394
 QY 761 CysLeuSerAspLeuValHisProG|uAsnAsnAlaThrAspValA|gThrTyrThrThr 780
 Db 2395 TGTATCTGATGATTTGGTTCATCCAGAAACAAATGCTACTGATGTCAAGAACTTACACACA 2454
 QY 781 ArgTyrGlnValPheTyrSerLySerG|yTrrpLysIleThrGlnG|uSerValLeuAla 800
 Db 2455 AGATACGAAGTTTCTGGTCCAAAGTCAGGGTGGAAATCACTGAAAGGCTCTGTTCTTGCA 2514
 QY 801 Ser 801
 Db 2515 TCA 2517

RESULT 5

US-10-600-070-9
 ; Sequence 9, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitna, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hongbo
 ; TITLE OF INVENTION: Plaseid Division and Related Genes and Proteins, and Methods of
 ; TITLE OF INVENTION: Use
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 2406
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-600-070-9

Alignment Scores:

Pred. No.: 0 Length: 2406
 Score: 4049.00 Matches: 799
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 2
 Query Match: 99.7% Indels: 0
 DB: 7 Gaps: 0

US-10-600-070b-2 (1-801) x US-10-600-070-9 (1-2406)

QY 1 MetG|uAlaLeuSerHisValG|yIleG|yLeuSerProPheGlnLeuCyArgLeuPro 20
 Db 1 ATGGAACCTCTGATCAGTCAGTCGCGATTCGTCCTCCCATTCCTCAATTATGCGCATTTACA 60
 QY 21 ProAlaThrThrLysLeuArgArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
 Db 61 CCGCGCAGCAACAAAGCTCCGACGTAAGCCACAAACACTCTCAACTATCTGCTCGCGCAGC 120
 QY 41 LysThrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSerSer 60
 Db 121 AAATGGGCGCAGCTCTCTCTCTCGACTTCAATTCCTCCGATTCCTCTCTCTCTCTCTCTCC 180
 QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
 Db 181 TTGGCCACCGCGCACCAACACCGCACTCTGTCCTCTGCCACCATATTTGAATCGTCCC 240
 QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlnValGlnThrHisPhe 100
 Db 241 GAAAGCCACGTCCTCCATCCCATGATTTCTTACCAAGATATTAGAGCTCAAAACACATTC 300

QY 101 LeuThrAspG|yIleArgArgAlaPheG|uAlaArgValSerLySerProGlnPheG|y 120
 Db 301 TTAAACCATGGAAATCAAAAGACATTTGGAAGCTAGGCTTTTGAACCCGCCCATTTGGCT 360
 QY 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCysG|uThrLeu 140
 Db 361 TTCAGCAGCAGCGCTTAAATCAGCCGAGACAGATTCCTCAAGCTGCTCGGAAACTCTG 420
 QY 141 SerAsnProArgSerArgArgG|uTyrAsnG|uG|yLeuLeuAspAspG|uGlnAlaThr 160
 Db 421 TCTAATCTCGCTCTTAAAGAGAGTACAAAGAGCTTCTTGTATGATGAAAGGCTACA 480
 QY 161 ValIleThrAspValProTrrpAspLyValProG|yAlaLeuCyValLeuGlnG|uG|y 180
 Db 481 GTATCATCTGATGTTCTTGGGATTAAGTTCCTGGGCTCTCTGTGATTCGAAGAGCT 540
 QY 181 G|yG|uThrG|uIleValLeuArgValG|yG|uAlaLeuLeuLyG|uArgLeuProLy 200
 Db 541 GGTAGACATGAGATGATTTCTTGGGTTGGTGAAGGCTCTGCTTAAAGAGAGGTTGCTAAG 600
 QY 201 SerPheLySerIleAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAsp 220
 Db 601 TCGTTTAAAGCAAGATGGTGTATATGATGCGCTGGCGTTCGATGTCGAGAGGAT 660
 QY 221 AlMetAlaLeuAspProProAspPheIleThrG|yTyrG|uPheValG|uGlnAlaLeu 240
 Db 661 GCTATGGCATTTGATCCACTGATTTTATTAATCTGTTATGAGTTTGTGAGAAAGCTTTG 720
 QY 241 LysLeuLeuGlnGlnGlnG|uAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAsp 260
 Db 721 AAGCTTTTACAGAGGAGAAAGAGACAGTACGCTTGCACCGGATTTACGTGCAAAATTTGAT 780
 QY 261 G|uThrLeuGlnG|uIleThrProArgTyrValLeuGlnLeuLeuG|yLeuProLeuG|y 280
 Db 781 GAGACTTTGAAAGAGACACTCGCGGTATGCTTGAGGACTACTTGCGTTACCGCTTGT 840
 QY 281 AspAspTyrAlaAlaLysArgLeuAsnG|yLeuSerG|yValArgAsnIleLeuTrrpSer 300
 Db 841 GATGATTAACGCTCGGAAACAACTAAATGTTTAAAGCGTGTGCGGATATTTTGCGTCT 900
 QY 301 ValG|yG|yG|yAlaSerAlaLeuValG|yG|yLeuThrArgG|uLySphMetAsn 320
 Db 901 GTTGAGAGAGTGGAGATCAGCTCTTGTGGGGGTTGACCGGTGAAGATTATGAAT 960
 QY 321 G|uAlaPheLeuArgMetThrAlaAlaGlnGlnValAspLeuPheValAlaThrProSer 340
 Db 961 GAGCGCTTTTATGAATGACAGCTGCTGACAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
 QY 341 AsnIleProAlaGlnSerPheG|uValTyrG|uValAlaLeuAlaLeuValAlaGlnAla 360
 Db 1021 AATATTCACAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGGCTCAAGCT 1080
 QY 361 PheIleG|yLysProHisLeuLeuGlnAspAlaAspLySerGlnPheGlnGlnLeuGln 380
 Db 1081 TTTATTGTTAAGAGCACACTTTTACAGAGTGTGATAGCAATTCAGCAACTTCAG 1140
 QY 381 GlnAlaLysValMetAlaMetG|uIleProAlaMetLeuTyrAspThrArgAsnAsnTrrp 400
 Db 1141 CAGGCTTAAGGTAAATGGCTATGAGATTCCTGCGCATGTTGTATATACACGAAATTAATGG 1200
 QY 401 GlnIleAspPheG|yLeuGlnArgG|yLeuCyAlaLeuLeuIleG|yLysValAspG|u 420
 Db 1201 GAGATAGACTTCGCTTGAAGAGGAGCTTGTGACCTGCTTATAGGCAAAAGTTGATGAA 1260
 QY 421 CysArgMetTrrpLeuG|yLeuAspSerG|uAspSerGlnTyrArgAsnProAlaIleVal 440
 Db 1261 TGCCTATGATGTTGGGCTTAAACAGAGGATTCACATATATGAAATCAGCATATTGG 1320
 QY 441 GluPheValLeuGlnAsnSerAsnArgAspAsnAspAspLeuProG|yLeuCyArg 460
 Db 1321 GAGTTTGTTTTGAAGATTCMAATCGTATGACAAATGATATCTCCCTGACATATGCMAA 1380
 QY 461 LeuLeuGlnThrTrrpLeuAlaG|yValAlaPheProArgPheArgAspThrLysAspLy 480

1381 TTGTTGAAACCTGGTTGGCAGGGGTTGTTCTTCTTCTAGTTCCAGAGCACCAAGATAA 1440
481 LysPheLeuLeuGlyAspTyrTyrAspAspPromeValLeuSerTyrLeuGluArgVal 500
1441 AAATTTAACTCGGGGACTACTAGATGATCTTATGTTGAGTTACTTGGAAAGAGTG 1500
501 GluValValGlnGlySerProLeuAlaAlaAlaThrMetAlaArgGlyLeuValGlu 520
1501 GAGGTAGTCAAGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1560
521 HisValValAlaSerAlaMetGlnAlaLeuGlnValPheProSerArgTyrThrAsp 540
1561 CATGTGAAGCTAGTGTCTATGACGAGCATCGAGAAATTTTCTTCTTCTTCTTCTTCTT 1620
541 ArgAsnSerAlaGluProGlyAspValGlnGluThrValPheSerValAspProValGly 560
1621 AGAAACTCGGCTGAAACCAAGAGATGTCAGAGACAGTGTATGATGATCTTCTTCTTCTT 1680
561 AsnAsnValGlyArgAspGlyGluProGlyValPheThrLeuAlaValArgProSer 580
1681 AACATGTAGCCCTGATAGGAGCTGTGTCTTATGCAAGAGCTGTAGAGCCCTCT 1740
581 GluAsnPheGluThrAsnAspTyrAlaAlaArgAlaGlyValSerGluSerSerValAsp 600
1741 GAAAACTTGAACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
601 GluThrThrValGluMetSerValAlaAspMetLeuValSerValValSerValLeu 620
1801 GAAACTACTGTGAAAGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
621 AlaAlaGlyValAlaAlaGlyLeuLeuSerLeuPheSerGlnLysTyrPheLeuLysSer 640
1861 GCTGCTGGTGGGCAATTGGACTGATTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
1921 AGCTCATCTTTCACGCAAGATATGTTCTTCTTATGATGATGATGATGATGATGATGAT 1980
661 GlySerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGlu 680
1981 GGGTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATGATG 2040
681 AsnIleValSerLysTyrGlnLysIleValSerLeuAlaPheGlyProAspPheArgIle 700
2041 AATATGATTCAGAGTGCAGAGATTAAGTCTGCTGCTTGGGCTTGTATCACCAGATA 2100
701 GluMetLeuProGluValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAla 720
2101 GAAATGTTACCAAGGTTTGGATGGGCGAATGCTGAAGATTGGACTGACAGAGCAGCT 2160
721 GluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuLysLeuSerValAspSer 740
2161 GAAACTCGGCAAGCTTGGTGTATGATTAACCTGTTGAAACATATCTGTTGACAGT 2220
741 ValThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGlnGluSerAla 760
2221 GTGACACTCTCAGCAGATGGAAACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
761 CysLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrThr 780
2281 TGTCTATCTGATTTGGTTCAATCCAGAAACATCTGATGTCAGAACTTACACACACA 2340
781 ArgTyrGluValPheTyrSerLysSerGlyTyrLysIleThrGlnGlySerValLeuAla 800
2341 AGATACGAAAGTTTCTGCTCAGAGTCAAGGTTGAAATCACTGAAGCTCTGTTCTTGA 2400
801 Ser 801
2401 TCA 2403

RESULT 6
US-10-600-070-3

Sequence 3, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Oseeryoung, Katherine W.
APPLICANT: Vichia, Stanislav
APPLICANT: Kosharova, Olga A.
APPLICANT: Gao, Hong
TITLE OF INVENTION: Placoid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 3667
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-600-070-3
Alignment Scores:
Pred. No.: 0 Length: 3667
Score: 3837.50 Matches: 800
Percent Similarity: 78.9% Conservative: 0
Best Local Similarity: 78.9% Mismatches: 1
Query Match: 94.4% Indels: 214
DB: 7 Gaps: 5
US-10-600-070B-2 (1-801) x US-10-600-070-3 (1-3667)
QY 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGlnLeuCyArgLeuPro 20
DB 481 ATGAAAGCTTGATGATCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 21 ProAlaThrThrIleValArgArgSerHisLeuThrSerThrThrIleCysSerAlaSer 40
DB 541 CCGGCGACACCAAGGCTCGACGTCACCAACCACTTACACATCTGCTGCTGCTGCTGCTGCT 600
QY 41 LysThrAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSer 60
DB 601 AAATGGGCGGACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
QY 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
DB 661 TTGCGACCGGCGACACCAACCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyValAlaGlnThrHisPhe 100
DB 721 GAAAGCGACGTCCTCCATCCCATGATTTCTTACAGGTATTAAGAGCTCAAAACACATTC 780
QY 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysPheProProGlnPheGly 120
DB 781 TTAAACGATGGAATTCAGAGAGCATTCGAAAGCTGTTGAAACCGCGCAATTCGCT 840
QY 121 PheSerAspAspAlaLeuLysSerArgArgGlnIleLeuGlnAlaAlaCysGlyThrLeu 140
DB 841 TTCAGGAGCAGAGCTTAAATCAGCCGAGACAGATTTCTTCAAGCTCTTGGCAAACTCTG 900
QY 141 SerAsnProArgSerArgArgGlyTyrArgGlnGlyLeuLeuAspAspGlyValAlaThr 160
DB 901 TCTTAATCTCTGCTCTGAAAGAGTCAATGAAGGCTCTTGTGATGATGAAGAGCTACA 960
QY 161 ValIleThrAspValProThrAsp 168
DB 961 GTATCATCGATGTTCTTGGGATTAAGTAAATTCATTTGCGAATTAATAACTTCTTC 1020
QY 169 -----LysValProGlyAla 173
DB 1021 GTTTTAATTTCAATGGAATGGAATAAGAGAACTTTTATCTAGTGAAGGTTCTCGGGGC 1080
QY 173 AluCyValLeuGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 193
DB 1081 TCTCTGTGATTCAGAAAGTGTGAGACTGAGATGATGATGATGATGATGATGATGATGATGAT 1140

QY 193 uleuylglnuargleupProlysserPheylglnaspValValleuValMetAlaLeuAl 213
 DB 1141 GCTTAAGAGAGGTGCTTACGTATGCTTTAGCAGAGAGAGTGTGTATGATGGGCTTGC 1200
 QY 213 aPheLeuaspValSerargaspAlaMetAlaLeuaspProPaspheilethrglyTy 233
 DB 1201 GTTCTCGATGTCTCGAGGGATGTCTATGGCATTTGATCCATTTTATTAATGAGTTA 1260
 QY 233 rglupheValgluValalauleylleuLeuGln----- 244
 DB 1261 TGAGTTGTGTAGAGAACCTTTGAAGCTTTTACAGGTAGTTGACTTGTGTAAATTTG 1320
 QY 244 ----- 244
 DB 1321 ACAGCCTGGCTTATTAAGAACTTTCTGATTTGATTACTTTGTTATTAAGCTTGTGTA 1380
 QY 245 -gluGluGlyAlaserSerleuAlaProaspLeuargAlaGlnileaspGluThrLeuG 264
 DB 1381 GGAGGAAGAGCAAGTAGCTTGCACCGAATTAAGTGCAAAATTAATGATGAGACTTTGGA 1440
 QY 264 uGluileThrProargTyValleuGluLeuLeuGlyLeuPProleuGlyaspAspTyAl 284
 DB 1441 AGAGATCACTCCGCTTATGTCTTGAGCTACTTGCTTACCGCTTGCTGATGTATGCGC 1500
 QY 284 aAlalyargleuaspnglyleuSerGlyValargAsnilleuThrSerValGlyGly 304
 DB 1501 TGGAAGAAAGCTAAATGTATTAAGCGGTGTGCGGAATATTTGTGTGTGTGAGAGAG 1560
 QY 304 yGlyAlaserAlaLeuValGlyGlyLeuThrArgGlylyPheMetAsnGlnAlaPhele 324
 DB 1561 TGAGAGATCAAGCTTCTTGTGGGGTTTGACCCCGGAGAAAGTTATGAATGAGGGCTTTT 1620
 QY 324 uArgMetThrAlaAlaGluGlnVal----- 332
 DB 1621 ACCAATGACAGCTGTGAGCAGAGT-ATACAGTTTACATCTTTTAAATTTCTTACG 1679
 QY 333 ----- 336
 DB 1680 ATGATATTAACCTTAAGCTTTCTCATTTAAATGATGTGTGTGAGGTATCTTTTGT 1739
 QY 336 AlaThrProSerAsnilleProAlaGluSerPheGluValTyArgValAlaLeuAlaLe 356
 DB 1740 AGCTACCCCAAGCAATTTCCAGCAGAGCTCATTTGAAGCTTACAAATTTGCACTTGCCT 1799
 QY 356 uValAlaGlnAlaPheileGlylylylyProHileuLeuGlnAspAlaAspLyGlnPh 376
 DB 1800 TGTGGCTCAAGCTTTTATGTGTAAAGAACCAACCTTTTACAGGATCTGATTAAGCAAT 1859
 QY 376 eGlnGlnLeuGlnAlaAlaValMetAlaMetGlnileProAlaMetLeuTyArgAspTh 396
 DB 1860 CCAAGCAACTTCACAGAGCTTAAGGTAAATGTGTGAAATTCCTGCGATGTGTATGATAC 1919
 QY 396 rArgAsnAsnTrpGlnileaspPheGlyLeuGluArgGlyLeuCyAlaLeuLeuileG 416
 DB 1920 ACCGAATTAATGGAGATAGACTTCCGTCTAGAAAGGGAGACTCTGTCACTGCTTAAGG 1979
 QY 416 yLyValaspGlyCyaspArgMetTrpLeuGlyLeuaspSerGlyaspSerGlnTyArgAs 436
 DB 1980 CAAGTGTGATGAATGCCGTATGTGCTTGAGGCTTAAGACATGAGGATTCACAAATATAGAA 2039
 QY 436 nProAlaileValGluPheValleuGluAsnSerAsnArgaspAspAspAspLeuP 456
 DB 2040 TCCAGCTATGTGAGATTTTGTGTGAGAAATTCAAATCGTATGATCAATATATATCTCCC 2099
 QY 456 oglyLeuCyAlaLeuLeuGluThrTrpLeuAlaGlyValAlaPheProArgPheArgAs 476
 DB 2100 TGAACATATGCAATTTGTGAAACCTTGCTTGCAGGGGCTTCTCTTCTAGCTTACAGA 2159
 QY 476 pThrlyaspLylyaspPheleuGlyaspTyTyArgaspProMetValleuSerTy 496
 DB 2160 CACCAAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCTATGTGTTTGAAGTTA 2219

QY 496 rLeuGluArgValGluValAlaGlnGlySerProLeuAlaAlaAlaThrMetAlaAr 516
 DB 2220 CTTGGAAAAGAGTGAAGTATGTTCAAGGTTCTCTTACGTCTGCTCAACTATAGCAG 2279
 QY 516 gileGlyAlaGluHileVallyAlaSerAlaMetGlnAlaLeuGlnlyValPheProSe 536
 DB 2280 GATTGAGCCGAGCATGTGAAGAGTACTAGTCTATGACAGGCACTGCAGAAATTTTCTTC 2339
 QY 536 rArgTyThrAspArgAsnSerAlaGluPProlyaspValGlnGluThrValPheSerVa 556
 DB 2340 CCGCTATACAGATTAAGAACTCGCTGAACCAAGAGATGTCCAGAGACAGGTATTAAGT 2399
 QY 556 lAspProValGlyAsnAsnValGlyArgaspGlyGluProGlyValPheileAlaGlyAl 576
 DB 2400 AGATCTCTGTGTGATACATATAGCCGTGATGTGAGCTCTGTCTTTATTTGAGAGAC 2459
 QY 576 aValArgProSerGlyuAsnPheGluThrAsnAspTyAlaileArgAlaGlyAlaSerG 596
 DB 2460 TGTAAAGACCTCTGAAACTTTGAACTAATGATTAATGCAATTCAGACTGGGGTCTCAGA 2519
 QY 596 uSerSerValaspGluThrThrValGluMetSerValAlaaspMetLeuylsglyAlaSe 616
 DB 2520 GAGTAGGGTGTATGAACCTACTGTGAATGTCCGTGTGATATGTTAAAGAGGCAAG 2579
 QY 616 rVallyleleuAlaAlaGlyValAlaileGlyLeuileSerleuPheSerGlnlyTy 636
 DB 2580 TGTGAAGATCTTACGTCTGTGTGTGCAATTTGACATATTTCACTGTTCAGCCGAAAGTA 2639
 QY 636 rPheLeuLySerSerSerSerPheGlnArgLyaspMetValSerSerMetGlySerAs 656
 DB 2640 TTTTCTTAAGACAGCTCATCTTTTCAACCAAGATATGTTCTTCTATGAAATCTGA 2699
 QY 656 pValAlaThrIle----- 660
 DB 2700 TGTGCTTACCATGATATGATTAATGATGCAATTTTCATATATCTGATTCCTCAAAAT 2759
 QY 660 ----- 660
 DB 2760 ATGCTGTGTTGTGAGCTAAGAACATAGTTCCCACTTAATACATGTCCTCAAAAGTTGAC 2819
 QY 660 ----- 660
 DB 2820 CAAGATTAACAAGTGTGTAGTAATTTCACTAATTAATGCTGCTGAAATTTTGTATCA 2879
 QY 660 ----- 660
 DB 2880 ACTGTAGACAGAAATGTAAATTTCACTTCACATTTCTGTTAGATTAAGATTA 2939
 QY 660 ----- 660
 DB 2940 GAGATTCCTTAAGTGTGCTTGTCCAACTTTTCTTGAATTTTCTTTTGCATT 2999
 QY 661 --GlySerValArgAlaaspAspSerGlyuAlaLeuProArgMetAspAlaArgThrArg 680
 DB 3000 TAGGGTCAAGTCAAGAGCTGAGATTTCAAGAACTTCCCAAAAGATGCTAAGAGCTCAG 3059
 QY 680 lAsnilleValSerlyeTrpGlnlylylelySerleuAlaPheGlyProaspHileArg 700
 DB 3060 AGAATATAGTATCCAAAGTGCAGAAAGATTAAATGCTGCTTGTGGGCTGATCACCGCA 3119
 QY 700 leGluMetLeuPProGlu----- 705
 DB 3120 TAGAAATGTATACAGAGAGGTGAGGAATTAATCTAATTCATCAATTCATTCATTCATTC 3179
 QY 706 -----ValLeuaspGly 709
 DB 3180 TTGACATATATATAGCTGTGCTGCTTGTGATTCGTATTAATTAATGATTTTGTGATGG 3239
 QY 710 ArgMetLeuLylyleThrPheAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTy 729
 DB 3240 CGAATGCTGAAGATTTGTGACTGACAGAGCAAGCTGAAATGCGGAGCTTGGGTGGTTAT 3299
 QY 730 AspTyThrleuLeuLylyleuSerValaspSerValThrValSerAlaaspGlyThrArg 749

Db 3300 GATTATACACTGTTGAAGACTATCTGTGACAGTGTGACAGTCTCAGACAGATGGAACCGT 3359
Qy 750 AAlenValGluValThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGlu 769
Db 3360 GCTCTGGTGAAGAACTCTGGAGAGCTGTCTGTCTATCTGATTTGTTTCATCCAGAA 3419
Qy 770 AspaAnaIaThrAspValArgThrTyrThrArgTyrGluValPheTyrSerLysSer 789
Db 3420 AACATCTACTGATGTCAAGAACTTACACAAAGATACGAAGTTTCTGTCTCAAGTCA 3479
Qy 790 GYTTrpValLeuThrGluGlySerValLeuAlaSer 801
Db 3480 GGGTGAAATCACTGAAGGCTGTCTTCTTCATCA 3515
RESULT 7
US-10-600-070-10
; Sequence 10, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislaw
; APPLICANT: Kosharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-10
Alignment Scores:
Pred. No.: 0 Length: 3667
Score: 3823.50 Matches: 798
Percent Similarity: 78.7% Conservative: 0
Best Local Similarity: 78.7% Mismatches: 3
Query Match: 94.1% Indels: 214
Gaps: 5
US-10-600-070b-2 (1-801) x US-10-600-070-10 (1-3667)
Qy 1 MetGluAlaLeuSerHisValGlyIleGlyLeuSerProPheGluLeuCybArgLeuPro 20
Db 481 ATGGAACTCTGAGTCAAGTCAGTCGGCATTTGGTCTCTCCCATTTCCAAATTATGCCATTACA 540
Qy 21 ProAlaThrThrLysLeuAlaGlyArgSerHisAsnThrSerThrThrIleCysSerAlaSer 40
Db 541 CCGGCGGCGCAAAAGCTCCGACGTAGGCACAAACCTCTTCACTATCTGTCTCGCAGC 600
Qy 41 LysTrpAlaAspArgLeuLeuSerAspPheAsnPheThrSerAspSerSerSerSerSer 60
Db 601 AAATGGCCCAAGCCTCTTCTCTCCGACTTCAATTTCACCTCCGATTCCTCTCTCTCC 660
Qy 61 PheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIleAspArgPro 80
Db 661 TTGCGCACCGCACCAACACGCGCAGCTCTGCTCTCTGCGCACCATCATTTGATGCGTCC 720
Qy 81 GluArgHisValProIleProIleAspPheTyrGlnValLeuGlyAlaGlnThrHisPhe 100
Db 721 GAACGCCACGCTCCCATCCCATTTGATTTTACCAAGATTTAGAGAGTCAAAACATTTTC 780
Qy 101 LeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGly 120
Db 781 TTAAACGATGGAATCAAGAGCACTTGCAGACTTGGGTTTCGAAACCGCGCAATTCGGT 840
Qy 121 PheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIaCysGluThrLeu 140

Db 841 TTCAGCGACGACGCTTTAATACCCGAGACAGATTTCTTCAAGCTGCTCGAAACTCTG 900
Qy 141 SerAsnProArgSerArgArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThr 160
Db 901 TCTAATCTCTGGCTCTAGAGAGAGTACATAGAAAGCTTCTTATATATGAAGAGCTACA 960
Qy 161 ValIleThrAspValProTTrpAsp 168
Db 961 GTCATCACTGATGCTTCCTGGATTAAGTAAATTTGATTCGGAATTAATAAGTTTCTTC 1020
Qy 169 -----LysValProGlyVal 173
Db 1021 GTTTTAATTTATCAAAATGGATTAAGAGAAAGAACTTTATCTAGTAAGAGTTCTGGAGC 1080
Qy 173 AleuCybValLeuGluGluGlyGlyGluThrGluIleValIleArgValGlyGluAlaLeu 193
Db 1081 TCTCTGTGATTTCCAGAGAGGTGTAGACTGATGATGTTCTTCGGGTGTGAGGCTCT 1140
Qy 193 ULeuValGluArgLeuProLysSerPheLysGlnAspValValLeuValMetAlaLeuAl 213
Db 1141 GCTTAAGAGAGGTTGGCTTAAGTCTTTAAGCAAGATGTGTGTTAGTTATGGCGCTTGC 1200
Qy 213 AspLeuAspValSerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyr 233
Db 1201 GTTTCGATGTCTCGAGGAGATGCTATGGCATTGATCCACCTGATTTTATCTGCTTA 1260
Qy 233 rGluPheValGluGluAlaLeuLysLeuGln 244
Db 1261 TGAGTTTGTGAGAAAGCTTTGAAGCTTTTACAGATAGTTGACTTCCTTGGTAATTGG 1320
Qy 244 ----- 244
Db 1321 ACAGCGTTGGCTTTAATAAGAACTTCTTGATTTGATTACTTTGTATAGCTCTTGATGA 1380
Qy 245 -GluGluGlyAlaSerSerLeuAlaProAspLeuAlaGlnIleAspGluThrLeuG 264
Db 1381 GGAGAAAGAGCAAGTAGAGCTTGACCGATTTACGTGACAAATTTAGATGAGCTTTGGA 1440
Qy 264 UGluIleThrProArgTyrValLeuGluLeuGluGlyLeuProLeuGlyAspAspTyrAl 284
Db 1441 AGAGATCATCCCGCTTATGTCTTGAAGCTACTTGGCTTACCGCTTGATGATATTAACG 1500
Qy 284 aaIaLysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTyrSerValGlyGly 304
Db 1501 TCGGAAAGACTAAATGTTTAAGCGGTGCGGAATATTTGTGCTGTGACGAGAG 1560
Qy 304 YGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLe 324
Db 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGTTTATGAAATGAGCGGTTTTT 1620
Qy 324 UArgMetThrAlaAlaGluGlnVal ----- 332
Db 1621 ATGAAATGACAGCTGTGAGCAGGT -ATACAGTTTATACCTTTTAAATTTCTTTAGC 1679
Qy 333 -----AspLeuPheVal 336
Db 1680 ATGATTAATTAAGTTTCTCATTTTAATGATGTTGTGTGAGGTTGATCTTTTGT 1739
Qy 336 LAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLe 356
Db 1740 AGCTACCCCAAGCAAAATTCACAGACAGTCAATTTGAAGATTGCAAGTTGCACCTGTCT 1799
Qy 356 vAlaIaGlnAlaPheIleGlyLysLysProHisLeuLeuGlnAspAlaAspLysGlnPh 376
Db 1800 TGTGGCTCAAGCTTTATTTGTTAGAGAGCCACCTTTTACAGAGATGCTGATTAAGCAATT 1859
Qy 376 eGlnGlnLeuGlnGlnAlaLysValMetAlaMetGluIleProAlaMetLeuTyrAspTh 396
Db 1860 CCAAGCAACTTCAACAGGCTTAAGTTAATGCTTAAGGATTCCTGCGATGTGTATGATAC 1919
Qy 396 rArgAsnAsnTrpGluIleAspPheGlyLeuGluValGlyGlyLeuCybAlaLeuLeuIleG 416
Db 1920 ACCGAATTAATTGGAGATAGACTTCTGCTTGAAGAAAGGAGACTCTGTGCACTGCTTATAG 1979

QY 416 YLYSVALASPGLUCYARGMETPLLEUGLYLEUASPSEGLUASPSEGLINTRYARGA 436
 DB 1980 CAAGATTGATGAAATGCCGTAATGCTGGCTTGAACAGTGAAGATTCAACATATAGAA 2039
 QY 436 nProalalileValaGluPheValleuGluuAenSerAsnArgAspAspAsnAspLeuP 456
 DB 2040 TCCAGCTATGTGAGATTGTTTGGAGAAATTCGATGATGACATGATGATGATCTCCC 2099
 QY 456 oGlyLeuCySlyLeuLeuGluThrTTPLeuAlaGlyValaValaPheProArgPheArg 476
 DB 2100 TGGACATATGCAAAATTTGTTGAAACCTGGTGGCAGGGGTTGTCTTCTGTTCAAGA 2159
 QY 476 pThrLysAspLysLysPheLysLeuGlyAspLysLysAspAspProMetValleuSerTy 496
 DB 2160 CACCAAGATATAAAATTTTAACTCGGAGACTGATGATGATCTATGGTTTGAATTA 2219
 QY 496 rLeuGluArgValaGluValaValaGlnGlySerProLeuAlaAlaAlaThrMetAla 516
 DB 2220 CTGGAAAGAGTGAAGTATGTCAGGGTTCTCTTATGCTGCTGCTGCACTATGGCAG 2279
 QY 516 gTleGlyAlaGluHisValaLysAlaSerAlaMetGlnAlaLeuGlnLysValaPhePro 536
 DB 2280 GATTGAGCCGACATGTGAACCTAGTCTATGCAAGCAGTGCAGAAAGTTTCTTCTTC 2339
 QY 536 rArgTyThrAspArgAsnSerAlaGluProLysAspValaGlnGluThrValaPheSerVa 556
 DB 2340 CCGCTATACAGATAGAACTCGGCTGACCCAGAGATGTCAGAGACAGTGTATAGTGT 2399
 QY 556 lAspProValaGlyAsnAsnValaGlyArgAspGlyGluProGlyValaPheIleAlaGlu 576
 DB 2400 AGATCTGTTGTAACATGATAGCGCTGATGCTGATGCTGCTTATTCAGAGAC 2459
 QY 576 aValaArgProSerGluAsnPheGluThrAsnAspTyrrAlaIleArgAlaGlyValaSerGI 596
 DB 2460 TGTAAACCTCTCGAAACCTTGAACCTAATGATGATGATGATGATGATGATGATGAT 2519
 QY 596 uSerSerValaAspGluThrThrValaGluMetSerValaAlaAspMetLeuLysGluLase 616
 DB 2520 GAGTACCGTGAAGAACTGATCTGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2579
 QY 616 rVallySileLeuAlaAlaGlyValaAlaIleGlyLeuIleSerLeuPheSerGlnTy 636
 DB 2580 TGTGAAGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2639
 QY 636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValaSerSerMetGluSerAs 656
 DB 2640 TTTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGAAATCTGA 2699
 QY 656 pValaIleThrIle----- 660
 DB 2700 TGTGCTACCATGATGATGATTAATGATGCAATTTTCATATATCTGCAATGCTCAAAAT 2759
 QY 660 ----- 660
 DB 2760 ATGCTTGTGTTGAGACTAAGAACATAGTTCCCACTTAATACATGTCCTCAAAAGTTGAC 2819
 QY 660 ----- 660
 DB 2820 CAAGATTACAAGTTGCTGAGTAATTTCACTAATATATGCTGCTGAATTTTGTATCA 2879
 QY 660 ----- 660
 DB 2880 ACTGTAAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTGAATAACGTAAGATTA 2939
 QY 660 ----- 660
 DB 2940 GAGATTGCTTATGATGCTTGTCCAACTTTCTTCTTGAATTTTCTTCTTGTGATTT 2999
 QY 661 --GlySerValaArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaG 680
 DB 3000 TAGGGTCAGTCAAGAGCTGAGATTCAGAAAGCACTTCCCAAGATGATGATGATGATGATGAT 3059

QY 680 luanIleValSerLysTrpGlnLysIleYsSerLeuAlaPheGlyProAspHisArgI 700
 DB 3060 AGAATATAGATTCAGATGTCAGAAAGATTAAGTCTGCTTTTGGCTTATCACCAGCA 3119
 QY 700 lGluMetLeuProGlu----- 705
 DB 3120 TAGAAATGTTACAGAGGTGAGGAATTAATCTACATTCATATCAATGTTGTGAAAACTG 3179
 QY 706 -----ValleuAspGly 709
 DB 3180 TTGACATGATTAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3239
 QY 710 ArgMetLeuLysIleTrpThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValaTy 729
 DB 3240 CGAATGCTGAAGTTTGGATGCTGACAGAGCTGAAATCTGCGAGCTTGGGTTGTTAT 3299
 QY 730 AspTyThrLeuLeuLysLeuSerValaAspSerValaThrValaSerAlaAspGlyThrArg 749
 DB 3300 GATTATACACTGTTGAACATATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCGT 3359
 QY 750 AlaLeuValaGluAlaThrLeuGluGluSerAlaCysLeuSerAspLeuValaHisProGlu 769
 DB 3360 GCTCTGTGGAACCACTGAGAGAGTGTCTGTCTATCTGATTTGTTGATCCAGAA 3419
 QY 770 AsnAsnAlaThrAspValaArgThrTyThrThrArgTyGluValaPheTrpSerLysSer 789
 DB 3420 AACATGCTATCTATGATGTCAGAACTTACACACAGATGATGATGATGATGATGATGAT 3479
 QY 790 GlyTrpLysIleThrGluGlySerValaLeuAlaSer 801
 DB 3480 GGGTGAATAATCACTGAAGGCTCTGTTCTTGATCA 3515
 RESULT 8
 US-10-600-070-126
 ; Sequence 126, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Oeteryoung, Katherine W.
 ; APPLICANT: Vitche, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hong
 ; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 126
 ; LENGTH: 2283
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-10-600-070-126
 Alignment Scores:
 Pred. No.: 1,14e-168 Length: 2283
 Score: 1775.50 Matches: 390
 Percent Similarity: 62.5% Conservative: 119
 Best Local Similarity: 47.9% Mismatches: 213
 Query Match: 43.7% Indels: 93
 DB: 7 Gaps: 16
 US-10-600-070B-2 (1-801) x US-10-600-070-126 (1-2283)
 QY 12 SerProPheGlnLeuCySArgLeuProAlaThrThrLysLeuArgArgSerHisAsn 31
 DB 40 GGCATATGCGCTTCCTCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 99
 QY 32 ThrSerThrThrIleCys---SerAlaSerLysTrpAlaAspArgLeuLeuSerAspPhe 50
 DB 100 CACCCCTCGCTGCTTC 159
 QY 51 AsnPhe-----ThrSerAspSerSerSerSerPheAlaThrAlaThr 65

```

Db      160 CACCTCTCCCAACCGCGCCCTCCGACCCGCGCTCCCGGCCCGCGCCGCGCC
      66 ThrThrAlaThrLeuValSerLeuProProSerIleAspArgProGluThrValPro 85
Db      220 GCGCCCTCCGCTCCCTTCCTGCTCCGCTTCCTCCGACCGCGCAACGCTCCCTCCG
      86 IleProIleAspPheTyrglnValIleGlnValIleThrIlePheLeuThrAspGlyIle 105
Db      280 CTCCAAGTCGATTTCTTCAAGGTTCTAGGGGACAGCCACATTTCTTGGCGATGGCATC
      106 ArgArgAlaPheGlnIleArgValSerLysProProGlnPheGlyPheSerAspAspAla 125
Db      340 AGAGAGCGCTTCAGGACAGGATAGCCAGCCAGCATGATGCTTACAGCAGGATGCTC
      126 LeuIleSerArgArgGlnIleLeuGlnAlaIleCysGlnIleThrLeuSerLeuProAspSer 145
Db      400 CTGTGTGCTGTGACAAATGCTGCAGATTGCCATGACACTCTCATGAAACCAAACTCC
      146 ArgArgGluTyrglnGlnGlyLeuLeuAspAspGlnGlnIleThrValIleThrAspVal 165
Db      460 CGCAGTCAGTATATGTCGCTTCCTTCTGAGAACCGTGAAGAAAGCTCTCACCAATGATAT
      166 ProTrpAspLysValProGlyAlaLeuCysValIleGlnGlnGlyGlyIleThrGluIle 185
Db      520 GCTGGGACAAAG-----GAGGCTGGGAGGACACTTGTCT
      186 ValLeuArgValGlyGlnAlaLeuLeuLysGlnArgLeuProLysSerPheLysGlnAsp 205
Db      553 GTGCTGTGAATCTGGAGAAAGCTTCTGGATCGGACCAAGCGCTTCAAGCAGAGAC
      206 ValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAsp 225
Db      613 GTGGTCTACCGCATGCTGCTGCTTATGTGATCTTCAAGGATGCTATGGACGACAC
      226 ProProAspPheIleThrGlyTyrglnPheValGlnGlnAlaLeuLysLeuGlnGln 245
Db      673 CCGCCAGATGATATGGCTGCTGCGAGGTGCTCGAAGGCGCTCTCAAGCTCTTGAGAA
      246 GlnGlyAlaLeuSerLeuAlaProAspLeuArgAlaGlnIleAspGlnIleThrLeuGlnGln 265
Db      733 GATGGAGCAAGCAATCTCGACCTGATCTGCTTCAAGATTGATGAATCTTCGAGAG
      266 IleThrProArgTyrglnValLeuGlnLeuLeuGlyLeuProLeuGlyAspAspTyraAla 285
Db      793 ATTACACCTGCTGTGTATGTGAGCTTCTCTCCCTTCTATGACACAGCATCATTAAG
      286 LysArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGly 305
Db      853 AAGCGCCAGAAAGGGCTTCAAGGTGCGAAGAAACATTTTGTGAGCGTTGCGAGGAGGT
      306 AlaSerAlaLeuValGlyGlyLeuThrArgGlnLysPheMetLeuGlnAlaPheLeuArg 325
Db      913 ATTTGACACCTGTGGAGGAGATTTTCTGTGAAGCCCTTCATGAACGAGGCTTTTGTAGG
      326 MetThrAlaIleGlnGlnValAspLeuPheValAlaThrProSerAsnIleProAlaGln 345
Db      973 ATGACATCAATTGAAGATGATTTCTTTCAAAAACCCGAATAGCATTTCTCTCGAA
      346 SerPheGlnValTyrglnValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLysLys 365
Db      1033 TGGTTTGAATTTTCAATGATAGCACTTGCAATGTGCTCAAGCAATTAATTAAGTAAAGG
      1033 TGGTTTGAATTTTCAATGATAGCACTTGCAATGTGCTCAAGCAATTAATTAAGTAAAGG 1092
Db      1093 CCACAAATTCATCATGATGAGCGGATGATCTTTTGAACACTCCAGAACTTCAACATAGGT
      366 ProIleLeuGlnIleAspValAspLysGlnPheGlnGlnLeuGlnGlnAlaLysValMet 385
Db      1153 TCT-----CATTAAGCTTATGATTAAT-----GAGATGAGACTTGGCA 1188
      386 AlaMetGlnIleProIleMetLeuTyrglnAspThrArgAsnAsnTrpGlnIleAspPheGly 405
      406 LeuGlnArgGlyLeuCysAlaLeuLeuIleGlyLysValAspGlnCysArgMetTrpLeu 425

```

```

Db      1189 TTGAAAGGCGATTCTGCTCATTTCTAGTCGAGATGTTTACGAATGCGAAATGTGGCTT
      426 GlyLeuAspSerGlnAspSerGlnTyrglnAsnProAlaIleValGlnPheValLeuGln 445
Db      1249 GGAATTTGAATTAAGAGCTTCAACCATTAAGAGACCCCAAAATTTCTAGATTATTTGAGCC
      446 AsnSerAsn--ArgAspAspAsnAspAspLeuProGlyLeuCysAlaLeuLeuGlnIle 464
Db      1309 AACTTACATCATGACGACAAAGATGATCTTCTTCAGGGCTGTGCAAGCTTTTGGAGACT
      465 TrpLeuAlaGlyValValPheProArgPheArgAspThrLysAspLysValPheLysLeu 484
Db      1369 TGGCTATCTTTGAGGTTTCTTCTAGACAGAGATCTCGGGCGCATGAGTTCAAGATT
      485 GlyAspTyrglnAspAspProMetValLeuSerTyrglnGlnArgValGlnValAlaGln 504
Db      1429 GAGATTAATTAAGATGATCAGAAAGTTTAAGCTACTTGAAGATGAGGAGGTGGTGT
      505 GlySerProLeuAlaAlaAlaIleThrMetAlaArgIleGlyValAlaGln----- 520
Db      1489 GCTTCTCATTTGGCTGCTGCTGCTGCTATGCAAACTTGCTCAACCTCAACCTGCA
      521 -----HisValLysAlaSerAlaMetGlnAlaLeuGlnLysValPhePro--SerArg 537
Db      1549 CTGTGACTGTGAATCAATGATCTATTCAGGCTTCAACAGGTTTTCATTGATAGAA
      538 TyrThrAspArgAsnSerAlaGlnProLysAspValGlnGlnIleThrValPheSerValAsp 557
Db      1609 CAGTTAGACAGTCAGCCATGGAATAACT-----
      558 ProValGlyAsnAsnValGlyArgAspGlyGlnProGlyValPheIleAlaGlnAlaVal 577
Db      1639 -----AAAGATGCG-----CCTGGGCGGATATCTT----- 1662
      578 ArgProSerGlnAsnPheGlnIleThrAsnAspTyraAlaIleArgAlaGlyValSerGlnSer 597
Db      1663 -----GAAATTTTACCG----- 1677
      598 SerValAspGlnIleThrValGlnMetSerValAlaAspMetLeuLysGlnAlaSerVal 617
Db      1678 -----GAAATGCAACCTGCTCATGATTCGAAGAAATGCGGCTTG 1716
      618 LysIleLeuAlaIleGlyValAlaIleGlyLeuIleSerLeuPheSerGlnTyrglnPhe 637
Db      1717 AAGATTATCTGCTGCGCCACACTGTGTGCACTGTGCGAGTATTTGGCGCCAAATAT-- 1773
      638 LeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGlnSerAsp--- 656
Db      1774 -----TTGCCTGTAAGAGGCCCTTTCTGCTATTTAGAGTGAAGCAT 1815
      657 -----ValAlaThrIleGlySerValArgAlaAspAspSerGlnAlaLeu----- 671
Db      1816 GGAATCTGTGGCACTTCTAATAGTGTGCACTGTCAAGATGATCTGCACTAGATGAAGAT 1875
      672 -----ProArgMetAspAlaArgThrAlaGlnAsnIleValSerLysTrpGln 687
Db      1876 CCAGTACATATTTCTGAATGATGATGCGAAGCTGCGAAGATATTTGTCGCAAGTGGCAG
      688 LysIleLysSerLeuAlaPheGlyProAspPheArgIleGlnMetLeuProGlnValLeu 707
Db      1936 AGTATCAAACTTAAGGCTTTGGGACCAAGAACATTCGGTGCATCATTTGCAAGAGTTCTT
      708 AspGlyArgMetLeuLysIleTrpThrAspArgAlaAlaGlnThrAlaGlnLeuGlyLeu 727
Db      1996 GATGGCAATGCTTAAGGTGTGATGATGACCGAGCAGCGAGATTAAGGTGATGGGTGG
      728 ValTyrglnTyrglnLeuLysSerValAspSerValThrValSerAlaAspGly 747
Db      2056 TTCTGGAGATATACATATCCGATGTGACGATGATGATGATCACTATCTCCCTAGATGTGT
      748 ThrArgAlaLeuValGlnAlaThrLeuGlnGlnSerAlaCysLeuSerAspLeuValHis 767
Db      2116 CGACGAGCGACTGTGAGGCTAGATGATGAGGCGAGGCCAACTTACTGATTTACTGAG 2175

```

```

Qy 768 ProGluAsnAspAlaThrAspValArgThrThrThrArgTyrGluValPheTrpSer 787
Db 2176 CCCAGAAACAAATGATTATATGACAAATAATACCTACCGGATAGATAGAGCTTCTCC 2235

Qy 788 Lys---SerGlyTyrPylIethrGluGlySerValIleuAlaSer 801
Db 2236 AACGTAGAGGGTGGAGATACGGAAGAGAGCACTCTCAAGTCG 2280

RESULT 9
US-10-424-599-129007
; Sequence 129007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129007
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT1847_87499C.1
US-10-424-599-129007

```

```

Alignment Scores:
Pred. No.: 3,27e-104 Length: 1146
Score: 1134.00 Matches: 223
Percent Similarity: 80.7% Conservative: 48
Best Local Similarity: 66.4% Mismatches: 53
Query Match: 27.9% Indels: 12
DB: Gaps: 3

```

US-10-600-070b-2 (1-801) x US-10-424-599-129007 (1-1146)

```

Qy 16 LeuGlyArgLeuProAlaThrThr-----LysLeuArgArg 28
Db 148 CTTCGACATCCGACCCACCACTCACTTCTTAACTTAACAATCTCTCCGT 207

Qy 29 Ser-----HisAnthrSerThrThrIleCysSerIaSerIyStrIaAspArgLeu 46
Db 208 TCCTCCCTCAGCGCGCGCGCTTCTTATCCGCCACCAATGCGGAGCGCTC 267

Qy 47 LeuSerAspPheAsnPhetrIserAspSerSerSerSerPheAlaThrAlaThrThr 66
Db 268 ATTGCCGATTTCCAAATTCCTCGGCGAGCGCGCGCTCC-----ACCTCCACCTCC 318

Qy 67 ThrAlaThrIleValSerLeuProProSerIleAspArgProGluArgHisValProIle 86
Db 319 ACTCTCAGCCCTCTCTCGCTCCCTCCCGCTCGATCTCTGAGCGCTAGCGTGCATC 378

Qy 87 ProIleAspPheTyrGluValIleuGlyAlaGlnThrHisPheLeuThrAspGlyIleArg 106
Db 379 CCGCTCGACCTGTAACCGCATTCCTCGGCGGAGACCGCATCTCTCGGCGAGCATTCGC 438

Qy 107 ArgAlaPheGluAlaArgValSerIyProProGluPheGlyPheSerAspAspAlaLeu 126
Db 439 AGGGCTTACAGAGCCAGATTCCTGAAAGCTCTCAGTACCCCTTCAGCAACGAGCTTTA 498

Qy 127 IleSerArgArgGlnIleuGlnAlaIaCysGlnThrIleuSerAsnProArgSerArg 146
Db 499 ATAGCGCGCGCAATCTCTCAGAGAGCGCTCGAAGACCTTAGCTATCTTACTTCACA 558

Qy 147 ArgGluTyrAsnGluGlyLeuLeuAspAspGluGluAlaThrValIleThrAspValPro 166
Db 559 AGAGATACAAATCAAGAGCTTGTGAGAGAGAGAGAGCGCATTTCTCACTCAATCTCT 618

```

```

Qy 167 TrpAspLysValProGluValIleuCysValIleuGlnGluGlyGlyGluThrGluIleVal 186
Db 619 TTGACAAATTTCTTGAGCGGTTGTGTGTGTGACAGAACTGAGAGAGAGAGCTGTG 678

Qy 187 LeuArgValGlyGluAlaLeuLeuGlyValArgLeuProLysSerPheLysGlnAspVal 206
Db 679 CTGAGATTTGGGAGGGTTTGCTTAGGAGAGGGTTCGAAAGAGCTTTAAGAGAGATGTT 738

Qy 207 ValIleValIleAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspPro 226
Db 739 GTGTGGCTATGGACACCGCATTTGTGAGGTGCAAGGAGATGCTATAGGCTGTGTCCA 798

Qy 227 ProAspPheIleThrGlyTyrGluPheValGluGluAlaLeuLysLeuGlnGluGlu 246
Db 799 CCGATTTCAATTGGCGGCTGTGAGATGCTGAGAGGGCGTTGAAAGCTTTTGACAGAAAGA 858

Qy 247 GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeuGluGluIle 266
Db 859 GGGGCAACAGCGCTAGCTCCAGATTTTACAGACCAATAATAGTAGAGCGCTTAGAAGATA 918

Qy 267 ThrProArgTyrValIleuGluLeuLeuGlyLeuProLeuGlyAspAspTyrAlaAlaLys 286
Db 919 ACCCCAGTTGTGTGTGGAACTTTAGCCTTGCTTGATGACGAACATCGAGCGCG 978

Qy 287 ArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGlyValAla 306
Db 979 AGGAGAGAGGTTCTTCTGTGTGTGCGTAACATTTGTGGCGGTTGTGTGAGAGGGGTGA 1038

Qy 307 SerAlaLeuValGlyGlyLeuThrArgGluLysPheMetAsnGluAlaPheLeuArgMet 326
Db 1039 GCAACAAATGCCCGGGGTTTCAACCGCGAAGACTTATGATAGGATCTTACACATG 1098

Qy 327 ThrAlaAlaGluGlnValAspLeuPheValAlaThrProSerAsnIle 342
Db 1099 ACAGCGGCTGAAACAGGTGAACTTTGTGTGACCAACCAAGTACTATT 1146

```

RESULT 10

US-10-425-115-81853

; Sequence 81853, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 81853

; LENGTH: 1411

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_174665C.1

US-10-425-115-81853

```

Alignment Scores:
Pred. No.: 8.53e-88 Length: 1411
Score: 973.00 Matches: 198
Percent Similarity: 69.8% Conservative: 54
Best Local Similarity: 54.8% Mismatches: 79
Query Match: 23.9% Indels: 30
DB: Gaps: 4

```

US-10-600-070b-2 (1-801) x US-10-425-115-81853 (1-1411)

```

Qy 12 SerProPheGlnLeuCysArgLeuProProAlaThrThrLysLeuArgArgSerHisAsn 31
Db 152 GCGCGCTTGGCTTCTTCAACGCGTGGCC-----CTCCGAGAGAGCCACAG 196

```

```

QY 32 Thr-----SerThrThrIleCysSerAlaSerIleThrAlaAspArgLeu 47
DB 197 CCGCCACCGCCCTGGCCCTCCACCTCCGCCGCCGCCACCGCCCTGCTTC 256
QY 48 SerAspPheAsnPhe-----ThrSerAspSerSer 57
DB 257 GCGGACTTCACACTCCTCCCGCCGCCGCCGCCGCCGCCGCCCTCTCTCC 316
QY 58 SerSerSerPheAlaThrAlaThrThrThrAlaThrLeuValSerLeuProProSerIle 77
DB 317 TCGTCCCGCTTC-----GTCGCCGATCTTCCCC 343
QY 78 AspArgProGluArgHisValProIleProIleAspPheTyArgIleValLeuGluAlaGln 97
DB 344 GAAGCCGCCGACCGCCCTGCTCCCGCTCCCGCTGCACTTCAAGATCTTGTCGGAG 403
QY 98 ThrHisPheLeuThrAspGlyIleArgArgAlaPheGluAlaArgValSerIleProPro 117
DB 404 CCACATTTCTAGCGGATGCGATTCGGAAGGCCGTCGAGTCGCGGATAGCCACCT 463
QY 118 GlnPheGlyPheSerAspAspAlaLeuIleSerArgArgGlnIleLeuGlnAlaAlaCys 137
DB 464 CAGTATGGGTACGACACAGAACTCTTGTGGCGACGCAATGCTGCAGATTGCCAT 523
QY 138 GlnThrLeuSerAsnProArgSerArgGluTyArgGluIleLeuAspAspGlu 157
DB 524 GATCTCTCAAAACAGAGCTCGCCACCGATACACCGTCGCTTCCGAGACCT 583
QY 158 GlnAlaThrValIleThrAspValProTrpAspIleValProGlyAlaLeuCysValLeu 177
DB 584 GATGCCGACCTACCATGATGATGCTGCTGCGATAGGATTCAGATGCTGCTGCT 643
QY 178 GlnGluGlyGlyIleThrGluIleValLeuArgValGlyIleValLeuLeuGluArg 197
DB 644 CAGAGAGCTGGGAGGACAACTGCTGTAGCACTGGAAGCACTTCTTCAGAGACCT 703
QY 198 LeuProIleSerPheIleGlnAspValIleLeuValMetAlaLeuAlaPheLeuAspVal 217
DB 704 CCACCTTAAGCGGCTTCAAGCAGATGCTGCTGCTGCAATGCGATTCGTTATGCGACATA 763
QY 218 SerArgAspAlaMetAlaLeuAspProProAspPheIleThrGlyTyArgIleValLeu 237
DB 764 TCAAGGGATGCTATGGCAGACAGCCCTCCAGATGATATCTGCTGTGAGGCTTGA 823
QY 238 GlnAlaLeuIleLeuLeuGlnGluGlyAlaSerSerIleLeuAlaProAspLeuArgAla 257
DB 824 AGGCACTGMAAGCTCTGCGAGAGGATGCGCAAGCAATCTTCACTGAACTGCTTCA 883
QY 258 GlnIleAspGluThrLeuGluGluIleThrProArgTyArgIleLeuGluLeuGluLeu 277
DB 884 CAGATTGATTAACCTTGGAGAGATTAACCTGCTGTGATTTGAGACTTCTTCCCTT 943
QY 278 ProLeuGlyAspArgTyArgAlaAlaIleArgLeuAsnGlyLeuSerGlyValArgAsnIle 297
DB 944 CCACTACGATGAATAAACAATAAATAAAGCCCAAGAGCTGCAAGGTGCAAAAAACATA 1003
QY 298 LeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGluLeu 317
DB 1004 TTGTGAGTGTGGCAGAGGCTGTGATGCTGCTGAGAGAGATTTCTCGTAGAGCC 1063
QY 318 PheMetAsnGluAlaPheLeuArgMetThrAlaAlaGluAlaLeuAspLeuPheValAla 337
DB 1064 TTATGATATAGGCTTCTTTCGACATACATCACTGACAGATGATTTCTTCTCTAAA 1123
QY 338 ThrProSerAsnIleProAlaGluSerPheGluValTyArgIleValAlaLeuVal 357
DB 1124 ACACCGAATAGCATACACCGAATGTTGATGATCTATATGTGCGACCTTGCCACATT 1183
QY 358 Ala 358
DB 1184 GCT 1186

```

```

RESULT 11
US-10-437-963-69933/c
; Sequence 69933, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbezuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 69933
; LENGTH: 1703
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7054C.1
US-10-437-963-69933

Alignment Scores:
Pred. No.: 2, 32e-86 Length: 1703
Score: 960.00 Matches: 227
Percent Similarity: 56.5% Conservative: 79
Best Local Similarity: 41.9% Mismatches: 147
Query Match: 23.6% Indels: 89
DB: 7 Gaps: 15

US-10-600-070b-2 (1-801) x US-10-437-963-69933 (1-1703)
QY 286 TyArgLeuAsnGlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGly 305
DB 1695 AGCGCCAGAGAGCGCTTCAAGGTGCGAGAAACATTTTGGAGCGTGGCAGAGAGGT 1636
QY 306 AlaSerAlaLeuValGlyGlyLeuThrArgGluLeuPheMetAsnGluAlaPheLeuArg 325
DB 1635 ATTGCTACCGCTGGAGAGAGATTTCTCGTGAAGCCTTATGAAACAGACCTTTTGGAG 1576
QY 326 MetThrAlaGluGluIleValAspLeuPheValAlaThrProSerAsnIleProAlaGlu 345
DB 1575 ATGACATCAATTGACAGATGATTTCTTTCAAAAAACCGAATAGCATTCCTCTGAA 1516
QY 346 SerPheGluValTyArgIleValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyLeu 365
DB 1515 TGGTTGAATTTAACAATGATGACATTCGACATGCTCAAGCAATTAATTAAGTAAGG 1456
QY 366 ProHisLeuLeuGlnAspAlaPheIleArgIleGlnIleGlnIleAlaValMet 385
DB 1455 CCACAATTCATCAATGATGGCGATGATCTTTTGAACAATCCAGAGTTCAACATAGGT 1396
QY 386 AlaMetGluIleProAlaMetLeuTyArgPheThrArgAsnAsnTrpGluIleAspPheGly 405
DB 1395 TCT-----CATTAAGCTTATGATTA-----GAGATGACCTTGCA 1360
QY 406 LeuGluArgGlyLeuCysValAlaLeuIleGlyLeuValAspGluCysArgMetTrpLeu 425
DB 1359 TTGGAAGGCGCATTCGCTCATTCATGATGCGAGATGTTAGCAAGTGCAGAAATGTGCTT 1300
QY 426 GlyLeuAspSerGluAspSerGlnTyArgAsnProAlaIleValGluPheValLeuGlu 445
DB 1299 GGAATTGATATATAGCTTCAACCATACAGAGACCGCGCAATATTAAGTTATGTAGCC 1240
QY 446 AsnSerAsn---ArgAspAspAsnAspAspLeuProGlyLeuCysIleLeuGluThr 464
DB 1239 AACTTAGCATGCAATGAGAGATGATTTCTTCCAGGGCTGTGCAAGCTTTTGGAGACT 1180

```

QY 465 TTPLeuAlaGlyValValPheProArgPheArgPheThrIlyAspIlyblyPheIlyLeu 484
 |||||
 DB 1179 TGGCTTATCTTTGAGCTTTTCTTGAAGAGCGGACACCGCGCGCGCGA----- 1129
 QY 485 GLYAspIlyTyThrAspProMetValLeuSerTyLeuGluArgValGluValGln 504
 |||||
 DB 1128 GCCCCGACCGGAGAGACCTTCCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1069
 QY 505 GLYSerProLeuAlaAlaAlaAlaThrMetAlaArgIleGlyAlaGluHis----- 521
 |||||
 DB 1068 GGAACCTCTAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1015
 QY 522 -----ValIyAlaSerAlaMetGlnAlaLeuGln 531
 |||||
 DB 1014 CCGCTCTCCGCGCTGCTCCACTTGTAAGTGAATGTAATGTAATGTAATGTAATGTA 955
 QY 532 LysValPheProSerArgTyThrAspArgAsnSerAlaGluProIlyAspValGlnI 551
 |||||
 DB 954 AAGGTTTTTCCATTGATAGAACAGTTAGACAGGTACAGC-----ATGGAAGA 907
 QY 551 UThrValPheSerValAspProValGlyAsnAsnValGlyArgAspIlyGluProIly 571
 |||||
 DB 906 TACT-----AAAGATGCG-----CCTGGGCG 886
 QY 571 LpheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAspIlyAlaIle 591
 |||||
 DB 885 ATATCTT-----GAAATTTTGACACG----- 864
 QY 591 GAlaGlyValSerGluSerSerValAspGluThrValGluMetSerValAlaAspMe 611
 |||||
 DB 863 -----GAAATATCACCTGCTCATGA 844
 QY 611 tLeuIyGluAlaSerValIlyIleLeuAlaGlyValAlaIleIlyLeuIleSerIe 631
 |||||
 DB 843 TTCGAAGAATCGCGCTTGAAGATTAATCTACTGGCGACCTGTTGCACTGTTGGAGT 784
 QY 631 uPheSerGluIlyTyThrLeuIySerSerSerPheGluIlyArgIlyAspMetValSe 651
 |||||
 DB 783 AATTGGGGCCAA-----GATTGCCACGTAAAGGCCCTTTG 745
 QY 651 rSerMetGluSerAsp-----ValAlaThrIleGlySerValArgIlyAlaAspSe 668
 |||||
 DB 744 CGCGCTTAGAGGAGCATGATCTGTGCGACGTTGTAATAGTGTGCACTCTACTGTA 685
 QY 668 rGluAlaLeu-----ProArgMetAspAlaArgThrAlaGluAs 681
 |||||
 DB 684 TCCTGCACTAGATGAAGATCCAGTACATATTCCTAGATGATGCGAAGCTGCGACAGA 625
 QY 681 nIleValSerIlyTyThrIlyIlySerIleuAlaPheGlyProAspHisArgIleG 701
 |||||
 DB 624 TATTGTTCCGAAATGCGACAGTATCAATCTTAAGGCTTGGACCGAACATTCGCTTGC 565
 QY 701 uMetLeuProGluValLeuAspGlyArgMetLeuIlyIleTyThrAspArgIlyAlaG 721
 |||||
 DB 564 ATCATTTGCAAGAGGATCTTGTATGCGCAACATGCTAAAGGTGTGAGCTGACGAGCG 505
 QY 721 uThrAlaGlnLeuGlyLeuValTyAspIlyTyThrLeuIlyLeuSerValAspSerVa 741
 |||||
 DB 504 GATTGACGTCATGGGTGCTTGGGATTAACATATCCGATGTCGATTAATAGCAT 445
 QY 741 lThrValSerAlaAspGlyTyThrArgAlaLeuValGluAlaThrLeuGluGluSerAla 761
 |||||
 DB 444 CACTATCTCTCGATGATGTCGACGAGGACTGTGAGGCTTACATTAATGAGCGGACCA 385
 QY 761 sLeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyThrThrAr 781
 |||||
 DB 384 ACTTACTGATGTTATGAGCCGCAAAACATGATGATATGACCAAAATACACTACCG 325
 QY 781 gTyTyGluValPheTyPheSerIlySerGlyTyTyPlyIleThrGluGlySerValLeuAl 800
 |||||
 DB 324 GATATGATGAGCTTCTCTCCAAAGGATGAGGATGATGATGATGATGATGATGATGAT 265
 QY 800 aser 801

DB 264 GTCG 261
 |||||
 RESULT 12
 US-10-425-115-57452
 ; Sequence 57452, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21 (53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425, 115
 ; NUMBER OF SEQ ID NOS: 2003-04-28
 ; SEQ ID NO 57452
 ; LENGTH: 1536
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_152391C.1
 ; US-10-425-115-57452
 Alignment Scores:
 Pred. No.: 6.6e-76 Length: 1536
 Score: 856.00 Matches: 191
 Percent Similarity: 58.2% Conservative: 82
 Best Local Similarity: 40.7% Mismatches: 126
 Query Match: 21.1% Indels: 70
 DB: 8 Gaps: 11
 US-10-600-070b-2 (1-801) x US-10-425-115-57452 (1-1536)
 QY 348 GluValTyGlyValAlaLeuAlaLeuValAlaGlnAlaPheIleGlyIlyblyPheProHis 367
 |||||
 DB 9 GAGATCTATAGTGTTCGCGCCACCGCTCGGCTCGACGATTCGCAAGTAAAGGCCCA 68
 QY 368 LeuLeuGlnAspAlaAspIlyGlnPheGlnLeuGlnAlaIlyAspIlyMetAlaMet 387
 |||||
 DB 69 TTCATCATGATGCGACATGTAATCTTCGAACCACTTCAGAAATTCATATATAGTTC 128
 QY 388 GluIleProAlaMetLeuTyThrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGlu 407
 |||||
 DB 129 -----TATCTTATGAAAC-----GAGATGACCTTCGCAATTAGAA 164
 QY 408 ArgGlyLeuCyAlaLeuLeuIleGlyIlyValAspGluCyAspMetTrpLeuGlyLeu 427
 |||||
 DB 165 AGGCACTCTGCTCGTACTTGTGGGATATTTGCAATTGCAATGTCAGATGCTTGGAA 224
 QY 428 AspSerGluAspSerGlnTyArgAsnProAlaIleValGluPheValLeuGluAsnSer 447
 |||||
 DB 225 GATTAAGAGTCCGTCGATATAGGAGACCCCAAAATATATGAGTTGTGGAATTA 284
 QY 448 Asn-----ArgAspAspAsnAspAspLeuProGlyLeuCyIlyLeuLeuGluIlyTrpLeu 466
 |||||
 DB 285 AGCATGAACGAAGATGATCTTCTTCAGGGCTGTGCAAGCTCTTGGAGACATGCTT 344
 QY 467 AlaGlyValValPheProArgPheArgAspThrIlyAspIlyblyAspPheIlyblyAsp 486
 |||||
 DB 345 GTTCCGAGGTATTTCCAGAGCAAGATACCCGAGGAGCGAGATTCAGATCGGAGAC 406
 QY 487 TyTyAspAspProMetValLeuSerTyLeuGluArgValGluValGlnIlySer 506
 |||||
 DB 405 TATTACAGAGACCCCAAAAGTTTAAGCTTAATGAAAGATGAGAGCGGCTGTGCTTC 464
 QY 507 ProLeuAlaAlaAlaAlaThrMetAlaArgIleGlyAlaGlu----- 520
 |||||
 DB 465 CATTATGTCGACGCTCCCTCCATAAGCAAAACTTGAGTGTGCTCACTGCGCTTGGT 524
 QY 521 HisValIlyAlaSerAlaMetGlnAlaLeuGlnIlyValPheProSerArgTyThrAsp 540

```

525 ACTGTAATAACAGTGTCTTCAAGCTTTCAGTAAGTTTCCCA-----569
541 ATGAASerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGly 560
570 -----TTGATAGAACAGCTTACGATGTCAGGC 596
561 AenAsnValGlyArgAspGlyGluProGlyValPheLeuAlaGluAlaValArgProSer 580
597 AAGGATATC-----CCAGT 611
581 GluAsnPheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAsp 600
612 GATGATCTTGATTAATCT-----CTTGAAAACCTCGCCCAAGAGAGTGTCT 659
601 GluThrThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu 620
660 GGAGACGCTATCCATGATTCAGAAATGACGCCCTTG-----AGATTATC 704
621 AlalaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSer 640
705 TCTGCTGGTACACTGCTTGCACTATTGTCAGTGTAGTCTCAAGTCTGCTCGCTAC 764
641 SerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIle 660
765 AAGTCA-----CTTCTGCTCTTAGGGCGCAGTATGGGTCTGTG 803
661 GlySerValArgAlaAspAspSerGluAla-----LeuProArg 673
804 GCGGTTGTTGACCTCATTTGATGTGTCAGCGCAGATGAGAGCCACTGAAATCCCTAGG 863
674 MetAspAlaArgThrIleGluAsnIleValSerLysTyrGlnLysIleLysSerLeuAla 693
864 ATGACACGCGAAGTTGGCTGAGACATTGTTCCGACGGTGGCAAGCATCAAGTCCAGGCT 923
924 TTGGGGCGCAACACATGTCACGGCATTCGAAAGAGTCTCGCGCGCAACATGCTGAG 983
714 IleTyrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeu 733
984 GTGTGGACGACCGACCGACGAGATGACGACCGACCGCTGCTTCTGAGTAGACGCGCTC 1043
734 LeuLysLeuSerValAspSerValThrValSerAlaAspGlyThrArgAlaLeuValGlu 753
1044 TCCGGCGCTGACGATGACGATCAACGCTCTCTCGTGACGCGCGCGCGGACGCGTGGAG 1103
754 AlaThrLeuGluGluSerAlaCysLeuSerAspLeuValHisProGluAsnAsnAlaThr 773
1104 GCGACGATCGACGAGGTGGCGCGCTTACGACGTCACCGACCTTAAGAACGACGACGCG 1163
774 AspValArgThrTyrThrArgTyrGluValPheTyrSerLys---SerGlyTyrLys 792
1164 TACGACACCAAGTACCGCGCGGTACGAGATGACTCTACGACGACGACGAGGTGAGG 1223
793 IleThrGluGlySerValLeuAlaSer 801
1224 ATCACCGAAGCGCAGTCTCAAGTGC 1250

```

RESULT 13
 US-10-600-070-132/c
 ; Sequence 132, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitsha, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Geo, Hongo
 ; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20

```

; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (127)..(127)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (520)..(520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-132

Alignment Scores:
Pred. No.: 2, 93e-69
Score: 784.00
Percent Similarity: 98.1%
Best Local Similarity: 98.1%
Query Match: 19.3%
DB: 7
Gaps: 0

US-10-600-070b-2 (1-801) x US-10-600-070-132 (1-561)

642 SerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIleGly 661
559 TCATCTTTTCAACGACGAAGATATGTTCTTCTATGAAACATGATCGCTACATAGGG 500
662 SerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGluAsn 681
499 TCAGTCAGAGCTGACCATTCAGAAACATTCCTCCAGATGATCTAGACTGACGAGAAAT 440
682 IleValSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHisArgIleGlu 701
439 ATAGTATTCAGAGTGGCGAAGATTAAGTCTCTGCTTGGGCTGATACCGCATAGAA 380
702 MetLeuProGluValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAlaGlu 721
379 ATGTTACAGAGGTTTGGATGGCGAATGCTGAAGATTGGACCTGACAGACGCTGAA 320
722 ThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLysLeuSerValAspSerVal 741
319 ACTGCCAGCTTGGGTGTTATGATTAACACTGTGAACATTACTGTTGACAGTGTG 260
742 ThrValSerAlaAspGlyThrArgAlaLeuValGluAlaThrLeuGluGluSerAlaCys 761
259 ACGTCTCAGCAATGGAACCCGTGCTGTGTGAAGCAACTGTGAGAGTCTGTCT 200
762 LeuSerAspLeuValHisProGluAsnAsnAlaThrAspValArgThrTyrThrArg 781
199 CTATCGATTGTGTTATTCAGAAACAAATGATCTAGATCTCAGAACTTACACACAGA 140
782 TyrGluValPheTyrSerLysSerGlyTyrLysIleThrGluGlySerValLeuAlaSer 801
139 TACGAAGTTTCNGGTCACAGTCAAGTGGGTAATAATCACTGAAGGCTCTGTTTGCATCA 80

```

RESULT 14
 US-10-600-070-184
 ; Sequence 184, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitsha, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Geo, Hongo
 ; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of

; Sequence 135, Application US/10600070

Db 600 GATGCTATGCGCTTGTCCCGCCAGATTTCATGTGTGCTGTGAGATGCTGGAAGGCCA 65

Wed Feb 22 13:40:19 2006

us-10-600-070b-2.p2n.rnpbm

Page 18

Search completed: February 21, 2006, 17:04:36
JOB time : 1617 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 21, 2006, 13:33:45 ; Search time 2594 Seconds
(Without alignments)
655.503 Million cell updates/sec

Title: US-10-600-070B-2
Perfect score: 4063
Sequence: 1 MEALSHVIGLSPFQCRLLP.....YEVFWKSGMKITGSLAS 801

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7204323 seqs, 1061406715 residues
Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=xlp
-Q=/abses/ABSSWEB.spool/US1060070/rnatc.21022006.112406.8805/app.query.fasta.1
-DB=Published Applications NA New -QFMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss03p
-USER=US1060070 @CEN 1 1 335 @rnatc.21022006.112406.8805 -NCPU=6 -ICPU=3
-NO MAMP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA New:
1: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
7: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:
12: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133.5	3.3	5880	12	US-11-000-688-226
2	121.5	3.0	2454	8	US-10-793-626-49
3	121.5	3.0	2454	8	US-10-793-626-1527
4	121.5	3.0	3295	8	US-10-793-626-3676

Result No.	Score	Query Match	Length	ID	Description
5	121.5	3.0	3742	8	US-10-793-626-4076
6	121.5	3.0	3985	8	US-10-793-626-3856
7	120	3.0	12507	12	US-11-136-527-2447
8	116	2.9	4965	8	US-10-485-517-43
9	114.5	2.8	163317	12	US-11-117-187-212
10	114	2.8	3327	8	US-10-793-626-3414
11	112	2.8	1908	11	US-11-098-686-9088
12	112	2.8	1457619	12	US-11-098-686-8739
13	111.5	2.7	11115	8	US-10-513-786-6
14	111.5	2.7	11115	8	US-10-513-786-8
15	111	2.7	3252	12	US-11-136-527-3210
16	110.5	2.7	3389	9	US-11-072-512-842
17	110	2.7	3002	9	US-11-031-206-183
18	109.5	2.7	1983	8	US-10-467-657-1265
19	109	2.7	1209	8	US-10-858-730-58
20	109	2.7	6914	12	US-11-000-688-1053
21	109	2.7	10302	12	US-11-127-832-23
22	108	2.7	2724	8	US-10-467-657-1069
23	108	2.7	4767	12	US-11-000-688-952
24	108	2.7	4767	12	US-11-169-041-99
25	108	2.7	8503	12	US-11-124-368A-45
26	108	2.7	8512	12	US-11-124-368A-46
27	107	2.6	2358	12	US-11-055-822-225
28	107	2.6	2358	12	US-11-055-822-713
29	107	2.6	2358	12	US-11-055-822-755
30	107	2.6	2810	12	US-11-168-476-1
31	107	2.6	3158	9	US-11-031-206-187
32	107	2.6	10300	8	US-10-947-249-122
33	106.5	2.6	3060	8	US-10-793-626-3495
34	106.5	2.6	9286	12	US-11-136-527-578
35	106.5	2.6	10705	12	US-11-136-527-2232
36	105.5	2.6	3828	7	US-10-724-598-48
37	105.5	2.6	13144	7	US-10-724-598-41
38	105	2.6	3931	12	US-11-136-527-3988
39	105	2.6	11736	12	US-11-000-463-218
40	104.5	2.6	8730	12	US-11-087-100-1
41	104.5	2.6	8730	12	US-11-087-084-1
42	104.5	2.6	8730	12	US-11-087-088-1
43	104	2.6	1881	8	US-10-467-657-5431
44	104	2.6	2238	8	US-10-858-730-262
45	104	2.6	6012	8	US-10-467-657-83

ALIGNMENTS

RESULT 1
US-11-000-688-226
; Sequence 226, Application US/11000688
; Publication No. US20050287544A1
GENERAL INFORMATION:
APPLICANT: BERTUCCI, Francois
APPLICANT: HOULGATTE, Remi
APPLICANT: BIRNBAUM, Daniel
TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
FILE REFERENCE: 1423-R-03
CURRENT APPLICATION NUMBER: US/11/000,688
CURRENT FILING DATE: 2004-12-01
PRIOR APPLICATION NUMBER: US 60/525,987
PRIOR FILING DATE: 2003-12-01
NUMBER OF SEQ ID NOS: 1596
SOFTWARE: PatentIn version 3.2
SEQ ID NO 226
LENGTH: 5880
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial sequences: primer
NAME/KEY: misc_feature
LOCATION: (1)..(5880)
OTHER INFORMATION: eukaryotic translation initiation factor 4
US-11-000-688-226


```

Qy 402 -----IleAspPheGlyLeu---GluArgGlyLeuCysAlaLeuLeuIleGlyVal 418
Db 1393 GGGGTTTACATCTGCTTATCTGAAGAAATATCCCTGAAGTAAGTGGT----- 1446
Qy 419 AspGluCysArgMetPheGlyLeu-----AspSerGluAspSer 432
Db 1447 -----TCGACAGGATATCTTTAACTAAATTAATGAAGTAATGATCAGAT 1491
Qy 433 GlnTyrArgGlnProAlaIleValGluPheValLeuGlnLysSerAlaArgAspAsn 452
Db 1492 CGTTTATTGAAAT-----CTTGAAAGATACACTTATTAACCTGCTATTCGACAAAC 1542
Qy 453 AspAspLeuProGlyLeuCysLysLeuLeuGlnThrTrpLeuAlaGlyValIlePhePro 472
Db 1543 GATGCTGTCAATTCATTAATTAAGCTGTTAAGAAAGCTGCTGCTCTT----- 1593
Qy 473 ArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr---TyrAspAspPro 491
Db 1594 -----AAAGATCCAAACGTCCAATCGGTAGTTTATTTCTTAGGACCT 1638
Qy 492 MetValLeuSerTyrLeuGlnArgValGluValIleGlnLysSerProLeuAlaIleAla 511
Db 1639 ACAAGTGTGGTAAACCTGAATGGCTGCTGCTTACCTGAATCTATGTTGCTGAAC 1698
Qy 512 AlaThrMetAlaArgIleGly---AlaGlnIleValLysAlaSerAlaMetGlnAlaLeu 530
Db 1699 GATGCAATGATTCCTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
Qy 531 GlnLysValIlePhePro-----SerArgTyrThrAspArgAsn 542
Db 1759 GTTGTCACCTCCAGATATGTAGACATGATGACGCGGTCAATGATGATGATGATGAT 1818
Qy 543 SerAlaGluProLysAspVal-----GlnGluThrVal 553
Db 1819 AGACGTAAACCATATCTGTGATTTATTTGATGAATTAAGAAAGCAATCCTGACGTA 1878
Qy 554 PheSerVal-----AspProValGlyAsnAsn 562
Db 1879 TTTAATTTCTTCTACAGTTTATGATGATGATGATGATGATGATGATGATGATGAT 1938
Qy 563 ValGlyArgAspGlyLysProGlyValPheIleAlaGluAlaValAlaArgProSerGluAsn 582
Db 1939 GTG-----GACTTCGTAATATCTGATGATATGATGATGATGATGATGATGATGAT 1992
Qy 583 PheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAspGluThr 602
Db 1993 TTACAGACCAACCGCTTGTGCTGTTTGGAGGTGCTTCAGAAAGGTAGT---GACTACGA 2049
Qy 603 ThrValGlnMetSerValAlaAspMetLeuLysGluAla----- 615
Db 2050 ACTGTCAAGAAACCAATGATGAAGAAATTAATAATTCATTCCGACCAATTCCTTAAC 2109
Qy 616 -----SerValLysIleLeuAla 621
Db 2110 CGTGTGATGATATTATGCTTCCACAACTTACAAAGATGATTAATAAGAAATTCCT 2169
Qy 622 AlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLys-----Tyr 636
Db 2170 ACAATGTGTAAATAACTTACTACCGCTTTCAGAGCAAAATTAATTAATTTGTTGTT 2229
Qy 636 rPheLeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
Db 2230 ACTGATTAAGCGAAGAAAGAAATTCGACAGACGATATGATCTGTAATATGCGCTAGA 2289
Qy 656 p 656
Db 2290 C 2290

```

RESULT 3
 US-10-793-626-1527
 ; Sequence 1527, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:

```

/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: P03480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ PRIOR FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1527
/ LENGTH: 2454
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1527

Alignment Scores:
Pred. No.: 0.0177 Length: 2454
Score: 121.50 Matches: 125
Percent Similarity: 33.0% Conservative: 93
Best Local Similarity: 18.9% Mismatches: 239
Query Match: 3.0% Indels: 204
DB: 8 Gaps: 28

US-10-600-070b-2 (1-801) x US-10-793-626-1527 (1-2454)

Qy 124 AspAlaLeuIleSerArgArgGlnIleLeuGlnAlaIleCysGluThrLeuSerAsnPro 143
Db 538 GATCCAGTGTGATGACGAGATGAATGAATTAATCTCGTGAATTAATGAAGTTTA----- 588
Qy 144 ArgSerArgArgGlnTyrAsnGlnGlyLeuLeuAspAspGlnLysAlaThrValIleThr 163
Db 589 ---AGTCGTGCTACTTAATAATATCTGCTGCTAATGCTGAA----- 627
Qy 164 AspValProTrpAspLysValProGlyAlaLeuCysValLeuGlnGlnGlyGlyGluThr 183
Db 628 -----CCCGGTGTT-----GGTAAACA 645
Qy 184 GlnIleValLeuArgValGlyGlnAlaLeuLeuLysGluArgLeuProLysSerPheLys 203
Db 646 GCAATGTCTGAAGGCTGTGGCAAGCAATGTTAAATGAATGATCCAGAACTTTAAAA 705
Qy 204 GlnAspValValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMetAla 223
Db 706 -----GACAAAGCTGTAATGTCA 723
Qy 224 LeuAspProAspPheIleThrGlyTyrGluPhe-----ValGluGlnAlaLeu 240
Db 724 TTGATATGGGTACAGTCGATGCTGACACTTAATATCTGATGATGATGATGATGATGAT 783
Qy 241 LysLeuLeuGlnGlnGlu----- 246
Db 784 AAAAAAGTATGAGAGAAATCCATCAAGCTGTAATGTTATTTATTCATGATGAACCTT 843
Qy 247 -----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThr 262
Db 844 CATACTTATGTTAGTGGCGCTGTGGCGCAGAAAGCAATGATGCA-----TCATATATT 897
Qy 263 LeuGlnGlnIleThrProArgTyrValLeuGlnLeuLeuGlyLeuProLeuGlyAspAsp 282
Db 898 TTTAAACCTGCTTATGCTGTGAGAAATTCATGATGATGATGATGATGATGATGATGAT 957
Qy 283 Tyr-----AlaAlaLysArg----- 287
Db 958 TATCGTAAATAATATGAAAAAGACGCTGATTAAGACGCTGTTTCAACCAATTCAAAGTG 1017
Qy 288 -----LeuAsnGlyLeuSerGlyValArgAsn 296
Db 1018 GATGAACCTACAGTTGAAGACACAGATTTGAATCTTTAAAGATTTACCTGACCGTTAGAG 1077
Qy 297 IleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyLysLeuThrArgGlu 316

```



```

Db      2743 TTAGATATGGGTACAGTCTGACGACATAAATATCTGTGTGAATTTGGAAGAAAGATTG 2684
Qy      241 LysLeuLeuGlnGlnGlu----- 246
Db      2683 AAAAAAGTTATGAGAAATCCATCAAGCTGTAAATCTTCTATTATTCATGAACTT 2624
Qy      247 -----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThr 262
Db      2623 CATACTTACTAGTGGCGCTGTGGCGGAGAGACCAATTGATGCA-----TCTAATATT 2570
Qy      263 LeuGlnGlnIleThrProArgTyrValLeuGlnLeuGlnGlyLeuProLeuGlnAspAsp 282
Db      2569 TTTAAACCTGCTTAGCTGTGAGAAATTCGAATGTATAGTCCACCAACATTAGATGAA 2510
Qy      283 Tyr-----AlaAlaLysArg----- 287
Db      2509 TATCGTAAATAATATAGAAAAAGCGCTGATAGAACGTGCTTTCAACCAATTCAAGTG 2450
Qy      288 -----LeuAsnGlyLeuSerGlyValArgAsn 236
Db      2449 GATGAACCTACAGTTGAAAGACAGATTGAAATCTTAAAGATTTACGTGACCGTTATGAG 2390
Qy      297 IleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyLeuThrArgGln 316
Db      2389 GCTCATCACAGAAATTATATCTCAGATGAGCTTTAGAACGGCGCTAAATTAGATGAT 2330
Qy      317 LysPheMetAsnGlnAlaPheLeuArgMetThrAlaIleGlnValAspLeuPheVal 336
Db      2329 CGGTATGTTTCAGATCGTTCTTGCCAGATTAAGCCATTGACTTAATGATGAGCGCAAT 2270
Qy      337 Ala-----ThrProSerAsnIleProAlaGluSerPheGln 348
Db      2269 TCAAAGCTTAGACTTAAAGATCATACACCGCAAGTAAATTTAAAGAGATTGAACAGAA 2210
Qy      349 ValTyrGlnValAla-----LeuAlaLeuValAlaGlnAlaPheIleGlyLys 364
Db      2209 ATTGATTAAGCTAAATAAAGAAAAAGATGCTGCGATGATGCTCAAGAAATTT----- 2159
Qy      365 LysProHisLeuLeuGlnAspAlaAspLysGlnPheGlnGlnLeuGlnAlaLysVal 384
Db      2158 ---GAAATATCCCGTAATTTAAGATTAAGCAATCTAAACCTGAAGCAATATGAA--- 2105
Qy      385 MetAlaMetGlnIleProAlaMetLeuTyrAspThrArgAsnEntPrgIu----- 401
Db      2104 -----GATCTAAATAATGAATGGAATAAATGACACA 2075
Qy      402 -----IleAspPheGlyLeu---GluArgGlyLeuCysAlaLeuLeuIleGlyLysVal 418
Db      2074 GGTGATTAGATATGCTTATCTGAAAGAAATATCTGTAAGTAATAGCTGGT----- 2021
Qy      419 AspGlnCysArgMetTTPLeuGlyLeu-----AspSerGluAspSer 432
Db      2020 -----TGACACGTAATTCCTTTAACTAAATTAATGAAGCTGAATTCAGAT 1976
Qy      433 GlnTyrArgAsnProAlaIleValGlnPheValLeuGlnLeuSerAsnArgAspAsn 452
Db      1975 CGTTATTTGAAT-----CTTGAAGATACACTTCAATAACGTGCATTGACAAAC 1925
Qy      453 AspAspLeuProGlyLeuCysLysLeuLeuGlnThrTrpLeuAlaGlyValAlaPhePro 472
Db      1924 GATGCTGTCAATTAATTAAGAAAGCTGTGAGAGAGCTGCTGCTGCTCT----- 1874
Qy      473 ArgPheArgAspThrLysAspLysLysPheLysLeuGlnAspTyr---TyrAspAspPro 491
Db      1873 -----AAAGATCCAAACGCTCAATCGGTAGTTTATTTTCTTAGACACT 1829
Qy      492 MetValLeuSerTyrLeuGluArgValGlnValGlnGlySerProLeuAlaAla 511
Db      1828 ACAGGTGGTAAACCTGAATTCGCTGCTGCTTACGTGATATCTATGTTGGGAAGAC 1769
Qy      512 AlaThrMetAlaArgIleGly---AlaGlnHisValLysAlaSerAlaMetGlnAlaLeu 530
Db      1768 GATGCAATGATTCGCTGATATATAGTGAATTTATGAGAAACATGCTGATGCAATTA 1709

```

```

Qy      531 GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
Db      1708 GTTGTGACACCTCCAGATATATGAGACATGATGACGCGGTCAATTGACGAAAAAGTT 1649
Qy      543 SerAlaGlnProLysAspVal-----GlnGluThrVal 553
Db      1648 AGACGTAAACCACTACTGTGATTTTATTTGATGAATAATGAGAAACACATCTCTGACGTA 1589
Qy      554 PheSerVal-----AspProValGlyAsnAsn 562
Db      1588 TTTAATATCTTCTTACAGTTTATGATGATGTCATTTAACAGATTAAGTTCGTACT 1529
Qy      563 ValGlyArgAspGlyLysProGlyValPheIleAlaGlnValAlaArgProSerGluAsn 582
Db      1528 GTC-----GACTTCGTAATACGTGATATATGATGATCTTAATGTGGAGCTCAAGAA 1475
Qy      583 PheGlnThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerSerValAspGluThr 602
Db      1474 TTACAGACCCACGCTTGTGTTGGAGGTCTCAGAAAGTATGAT---GACTACGAA 1418
Qy      603 ThrValGluMetSerValAlaAspMetLeuLysGlnAla----- 615
Db      1417 ACTGTCAAAAAACATGATGAAAGAAATTTAAATAATTCATCCGACCAAGATTTTAAAC 1358
Qy      616 -----SerValLysIleLeuAla 621
Db      1357 CGGTGATGACATTAATGCTTCCACAACTTACAAAGATGAATTTAAAGAAATTTGTT 1298
Qy      622 AlaGlyValAla-IleGlyLeuIleSerLeuPheSerGlnLys-----Tyr 636
Db      1297 ACAATGATGTAAATTAATCACTCAACCGCTTTCAGAGCAAAATTAATTAATTTGTTGTT 1238
Qy      636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
Db      1237 ACTGATTAACCGAAAGAAAAAATTGCAAGAAAGAAATGATATCTGAATATGGTCTAGA 1178
Qy      656 P 656
Db      1177 C 1177

RESULT 5
US-10-793-626-4076/c
; Sequence 4076, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OR INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4076
; LENGTH: 3742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4076

Alignment Scores:
Pred. No.: 0.0298 Length: 3742
Score: 121.50 Matches: 125
Percent Similarity: 33.0% Conservative: 93
Best Local Similarity: 18.9% Mismatches: 239
Query Match: 3.0% Indels: 204
DB: 8 Gaps: 28

US-10-600-070B-2 (1-801) x US-10-793-626-4076 (1-3742)

```

QY 124 AspaLeuIleuSerArgGlnIleuGlnIleuAlaCysGluThrLeuSerAsnPro 143
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1795 GATCAAGTCGTAAGACAGATTAAGAAATTAATCTGCTGTAATTAAGAGTTTA----- 1745
 QY 144 ArgSerArgArgGluTyrAsnGluGlyLeuLeuAspGluGluAlaThrValIleThr 163
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1744 ---AGTCGTCGTACTAAATAATATCCGCTTAATGGTGAA----- 1706
 QY 164 AspValProTrpAspIleValProGlyAlaLeuCyValLeuGlnGluGlyGluThr 183
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1705 -----CCCGGTGTT-----GGTAAACA 1688
 QY 184 GluIleValLeuArgValIleGluAlaLeuLeuGluArgLeuProLysSerPheLys 203
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1687 GCATTTCTGTAAGGCGCTTGGCGAGCATTTGTTAAATGAAGTACCAAGAACTTTAA 1628
 QY 204 GlnAspValIleValIleValIleMetAlaPheLeuAspValSerArgAspAlaMetAla 223
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1627 -----GACAAACGTGTAATGTCA 1610
 QY 224 LeuAspProProAspPheIleThrGlyTyrGluPhe-----ValGluGluAlaLeu 240
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1609 TTAGATTTGGGTACAGTCGTAGCTGCACTAAATATCGTGTGAATTTGAAGAAAGATTG 1550
 QY 241 LysLeuLeuGlnGluGlu----- 246
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1549 AAAAAGTTATGGAGAAATTCATCAAGCTGTGTAATGTTATTCTAATTTATCGATGAACCT 1490
 QY 247 -----GlyAlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThr 262
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1489 CATACTTACTGTTGGCGTGGTGGCGGAGAGAGCATTTGATGCA-----TCTAATATT 1436
 QY 263 LeuGluGluIleThrProArgTyrValLeuGluLeuLeuGlyLeuProLeuGlyAspAsp 282
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1435 TTAAACCTGCTTAGCTCGTGGAGAAATTCATGATGATAGTGCACAACTATTGATGAA 1376
 QY 283 Tyr-----AlaAlaLysArg----- 287
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1375 TATCGTAAATAATAGAAAAAGACGCTGCAATTGAAACGTCGTTTCAACCAATTCAAAGT 1316
 QY 288 -----LeuAsnGlyLeuSerGlyValArgAsn 296
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1315 GATGAACCTACAGTTGAAGACAGATGAAATCTTAAAGATTAACGTGACCTTAAGAG 1256
 QY 297 IleLeuTrpSerValGlyGlyGlyAlaSerAlaLeuValGlyGlyLeuThrArgGlu 316
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1255 GCTCATCACAGAAATTATCTCAGATGAAGCTTGAAGCGGCTGCTAAATTGAGTGAT 1196
 QY 317 LysPheMetAsnGlnAlaPheLeuArgMetThrAlaIleGluGlnValAspLeuPheVal 336
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1195 CGGTATGTTTCAGATCGTTCTTTCGAGATTAAGCCCATTAATTTGATGAAGCAAGT 1136
 QY 337 Ala-----ThrProSerAsnIleProIleGluSerPheGlu 348
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1135 TCAGAACTTGAAGTAAAGTCAATCAACCCCAAGTAAATTTAAAGAGATTGAACAAGAA 1076
 QY 349 ValTyrGluValAla-----LeuAlaLeuValAlaGlnAlaPheIleGlyLys 364
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1075 ATGATTAAGTAAATAAATAAGAAAGATGCTGAGATTGATGCTCAAGAAATTT----- 1025
 QY 365 LysProIleLeuLeuGlnAspAlaAspIleGlnPheGlnIleuGlnIleuGlnIleVal 384
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 1024 ---GAAATGCCCCCTAATTTAAGAGATTAACCAATCTTAAGAAAGCAATATGAA-- 971
 QY 385 MetAlaMetGluIleProAlaMetLeuTyrAspThrArgAsnAsnTrpGlu----- 401
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 970 -----GATCTTAAATTAATGAATGAAATGCAACA 941
 QY 402 -----IleAspPheGlyLeu-----GluArgGlyLeuCysAlaLeuLeuIleGlyLysVal 418
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 940 GGTGTTTATGATACCTGCTTATCTGAAGAAATATCTGCTGAAGTAATAGCTGGT----- 887

QY 419 AspGluCysArgMetTrpLeuGlyLeu-----AspSerGluAspSer 432
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 886 -----TGACAGAGTATTCCTTAACCTAAATTAATGAACGATTCAGAT 842
 QY 433 GlnTyrArgAsnProAlaIleValIlePheValLeuGluAsnSerAsnArgAspAsn 452
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 841 CGTTATTGAAAT-----CTGAAGATACACTTCATTAACGTCATTCGACAAAC 791
 QY 453 AspAspLeuProGlyLeuCyValLeuLeuGluThrTrpLeuAlaGlyValIlePhePro 472
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 790 GATGCTGTCAATTCATTAATTAAGCTGTGAAGAAAGCTCGTGGTGGTCTT----- 740
 QY 473 ArgPheArgAspThrLysAspIleLysPheLysLeuGlyLysAspTyr-----TyrAspAspPro 491
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 739 -----AAAGATCCAAACCGTCAATCGTATGTTTATTTTCTTAAGACCT 695
 QY 492 MetValLeuSerTyrLeuGlnArgValIleValIleGlnIleGlySerProLeuAlaIleAla 511
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 694 ACAGTGTGGGTAAACGTAATTTGGCTCGTCTTAAGCTGAATCTATGTTTGTGAAGAC 635
 QY 512 AlaThrMetAlaArgIleGly-----AlaGluHisValLysAlaSerAlaMetGlnAlaLeu 530
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 634 GATGCAATGATTCGCGTATGATGAGTAAATTTATGGAAGAACATGCTGTCAGTCGATTA 575
 QY 531 GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 574 GTTGTCACCTCCAGATATGTAGACATGATGACGCGGTCAATTGATCTGAAGAAAGTT 515
 QY 543 SerAlaGluProLysAspVal-----GlnGluThrVal 553
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 514 AGAGTAAACCATATCTGTGATTTTATGATGAATTTGAGAAAGACATCTGACGTA 455
 QY 554 PheSerVal-----AspProValGlyAsnAsn 562
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 454 TTAAATATTTCTTCAAGTTTATGATGATGTCATTTAACATTAAGTAAAGTCTGACT 395
 QY 563 ValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSerGluAsn 582
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 394 GTG-----GACTTCGCTAATCTGTGATTAATTAATGACTTAAATGCGAGCTCAAGAA 341
 QY 583 PheGluThrAsnAspTyrAlaIleArgAlaGlyValSerGluSerValAspGluThr 602
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 340 TTACAGACCAACGCTTGTGCTGTTTGAAGGTCTTCAAGAGTATGCT--GACTACGAA 284
 QY 603 ThrValGluMetSerValAlaAspMetLeuLysGluAla----- 615
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 283 ACTGTCAAGAAACCATGATGAAGAAATTAATAATTCATTCGACCAAGATTTCTTAAC 224
 QY 616 -----SerValLysIleLeuAla 621
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 223 CGTGTGATGACATTAATTTGCTTCCACAAACCTTAACAAAGATGAATTAAGAAATTTGTT 164
 QY 622 AlaGlyValAla-IleGlyLeuIleSerLeuPheSerGlnLys-----Tyr 636
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 163 ACATATATGTAATATAACTTACTCAACCGCTTTCAGAGCAAAATTAATATTGTTGTT 104
 QY 636 rPheLeuLysSerSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 103 ACTGATTAAGCGAAGAAATAATTTGCAAGAAAGATATGATCTGTAATATGTGTCTAGA 44
 QY 656 p 656
 DB 43 C 43

RESULT 6

US-10-793-626-3856/c
 ; Sequence 3856, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUI3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626

Db 962 CGTGTGATGACATTATGTTCTCCACAACTTACAAAGATGATTAAAGAAATGTT 903
 Qy 622 AlaglyValAla-1leglyLeuIleSerLeuPheSerGlnLys-----Tyr 636
 Db 902 ACAATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTGTT 843
 Qy 636 rphleuIleSerSerSerSerPheGlnArgLysAspMetValSerSerMetGluSerAs 656
 Db 842 ACTGATTAAGCGAAAGAAATTTGCGAAGAGATATGATCTGATATGTCGTCAGA 783
 Qy 656 p 656
 Db 782 c 782
 RESULT 7
 US-11-136-527-2447
 ; Sequence 2447, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; PRIOR FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2447
 ; LENGTH: 12507
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (12245)..(12245)
 ; OTHER INFORMATION: n is a, c, g, or t
 US-11-136-527-2447
 Alignment Scores:
 Pred. No.: 0.193 Length: 12507
 Score: 120.00 Matches: 146
 Percent Similarity: 32.7% Conservative: 98
 Best Local Similarity: 19.5% Mismatches: 313
 Query Match: 3.0% Indels: 190
 Db: 12 Gaps: 29
 US-10-600-070b-2 (1-801) x US-11-136-527-2447 (1-12507)
 Qy 53 ThSerAspSerSerSerSerPheAlaThrAlaThrThrAlaThrLeuValSer 72
 Db 1986 ACTCTTAAGCAAAAGCGGGGTGAAGAGACCGACCTGCCACCCGACGTCACAGCT 2045
 Qy 73 LeuProSerIleAspArgProGluArgHisValProIleProIleAspPheTyrGln 92
 Db 2046 GTTCCAGAGCTCCCAAGAGCGGGAGGCTGAGAACTGTCCCAAGCTTACTCTCAG 2105
 Qy 93 ValIleGlyAlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGluAlaArg 112
 Db 2106 GATCTGTCTGAGACCCACAAAGCTCAGTAT----- 2138
 Qy 113 ValSerIleProGlnIlePheIlePheSerAspAspAlaLeuIleSerArgArgGln--- 131
 Db 2139 -----ACAGGCTACTCTCCGATGCGCTCTCAGCTCCAGAGTGAG 2180
 Qy 132 -----1leuGlnAlaAlaCyGluThrLeuSer----- 141
 Db 2181 ATCACAGCTCTTATACAGCAAGAGCACTTACACAGTCCGGGGTGACAGGCCCA 2240
 Qy 142 -----AsnProArgSerArgArgGluTyrAsnGluIleLeuLeuAspArgGlu 158
 Db 2241 CGTCAACCCAGCCCTCAGAGCTCCACAAAGTGCGAGTAGTATGCGCCCTCCCTGAG 2300

Qy 159 Ala---ThrValIleThrAspValProTyrAspIleValProGlyAlaLeuCyValLeu 177
 Db 2301 GCCAGGCTGTGGCTCCAGTGTGAGTGAAGAGCAAGCCAGCCGAGCTGTGAA 2360
 Qy 178 GlnGluGlyGlyGluThrGluIleValLeuArgValGlyGluAlaLeu---LeuIleGlu 196
 Db 2361 GACCAGAAAGCGGCTCCCACTCCCTCCATCATGCCCCAGGCTTCGACTGAGAG 2420
 Qy 197 ArgLeuProIleSerPheIleArgIleAspValValLeuValMetAlaLeuAlaPheLeuAsp 216
 Db 2421 GAGCTGGGAGATATTCTAGAAAGATGACTCTCTGCTATGGG-----CGTCAG 2471
 Qy 217 ValSerArgAspAlaMetAlaLeuAspProProAspPheIleThr-----GlyTyrGlu 234
 Db 2472 AGGAGACAGCAAGACACAGCTGACTTTCAGATGACTTTGGACAGCCAGCTGAGCAGAC 2531
 Qy 235 PheValGluGluAlaLeuIleLeuIleGluGluGluIleValAspSerLeuAlaProAsp 254
 Db 2532 TACGTGAGAGACAGC-----AGTGAAGCGGCTGTCTCCCTCCACCTCAG 2579
 Qy 255 LeuArgAlaGlnIleAspGluThrLeuGluGluIleThrProArgTyrValLeuGluLeu 274
 Db 2580 CCCCAGCCAGGCGAGATGACTGATGAAAGATTCAGCGCGGAGATCTAGAAATG 2639
 Qy 275 LeuGlyLeuPro-----LeuGlyAspAspTyrAlaAlaIleAspArgLeuAsnGlyLeuSer 292
 Db 2640 AGCGCCGAGAAACAACTTGGAGAAAGACACACCGCTTCTCCGGCCAGGCTTGCC 2659
 Qy 293 GlyValArgAsnIleLeuTyrSerValGlyGlyGlyValAspAlaLeuValGlyGly 312
 Db 2700 -----AAACACGGTGCCAGAAAGCAAGTGCC----- 2726
 Qy 313 LeuThrArgGluIlePheMetLeuGluAlaPheLeuArgMetThrAlaAlaGluGlnVal 332
 Db 2727 AGACCCAGGCCCAAACTTACCAAGATCA----- 2756
 Qy 333 AspLeuPheValAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluVal 352
 Db 2757 -----GTACACTGCCCAAGAG----- 2774
 Qy 353 AlaLeuAlaLeuValAlaGlnAlaPheIleGlyIleValSerProHisLeuLeuGlnAla 372
 Db 2775 -----CGCTGCCCCAT-----AATGCC 2792
 Qy 373 AspIleGlnPheGlnGlnLeuGlnGlnAlaIleValValMetAlaMetGluIleProAlaMet 392
 Db 2793 ACCACAGTTACGAGAACTT-----CTGCTGAGAAAGCCCGACGACAG 2837
 Qy 393 LeuTyrAspThrArgAsnAsnTyrGluIleAspPheGlyLeuGluArgGlyLeuCyAla 412
 Db 2838 -----CCCACTGACGAGCCCTTCAAGGAGGATCTCCAGCGC 2873
 Qy 413 Leu---LeuIleGlyIleValAspGluCyAspMetTyr-----LeuGly 426
 Db 2874 TTTAAACATTTGTGATCAACACACACAGAGGATTTATGTCAACAGACTTAAGCTGGCG 2933
 Qy 427 LeuAspSerGluAspSerGlnTyrArgAsnProAlaIleValGluPheValLeuGluAsn 446
 Db 2934 CAAGGCTCTGATCCCAACCTAAGACCGGAGCTGAGCTGAGATGAGAGACCTTACAGCGC 2993
 Qy 447 SerAsnArgAsp----- 450
 Db 2994 TCGCTGAAAGACCGCTCTGCTGGGAGAGCACTCTTACACTGCCGCTCCAGCGCTAGC 3053
 Qy 451 -----AspAsnAspAspLeu 455
 Db 3054 TATACATCTGACACTCACCCACTCCCTGCTTCCCTGAGAGAGACAGTGAACAGACG 3113
 Qy 456 ProGlyLeuCyIleValLeuGluIleThrTrpLeuAla----- 467
 Db 3114 CCCAGCGGAGGAGCGCTGAGAAAGCAAGACAGCGGAGCCCGGACCGGCTCC 3173
 Qy 468 ---GlyValValPheProArgPheArgAspThrIleAspIleValSerPheIleGlyIleAsp 486


```

; LENGTH: 3327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3414

Alignment Scores:
Pred. No.: 0.169 Length: 3327
Score: 114.00 Matches: 126
Percent Similarity: 33.8% Conservative: 98
Best Local Similarity: 19.0% Mismatches: 214
Query Match: 2.8% Indels: 224
DB: Gaps: 31

US-10-600-070b-2 (1-801) x US-10-793-626-3414 (1-3327)
QY 71 ValSerLeuProSerIleAspArgProGluArgHisValProIleProIleAspPhe 90
DB 869 TTAACCTTAATCCCTCAGTTGAT-----ATATCTTATCCTTAGAT--- 910
QY 91 TyrGlnValLeuGlnAlaGlnThrHisPheLeuThrAspGlyIleArgArgAlaPheGlu 110
DB 911 -----CAGTTTAATTAGATACCTGTTAATAGGGA----- 940
QY 111 AlaArgValSerIleProProGlnPheGlyPheSerAspAspAlaLeuIleSerArgArg 130
DB 941 TCTCAACACAGCAAA----- 955
QY 131 GlnIleLeuGlnAlaIleAspGluThrLeuSerAsnProArgSerArgArgGluTyrAsn 150
DB 956 -----ACAGCAGCGCGTAAAGATTAAGTTACTAGATATTCTCGAATTTGCT 1006
QY 151 GluGlyLeu-----LeuAspAspGlu 157
DB 1007 GAAGACGTATATAGCAAGTGGTTTGTAGTGAGCATTAGGTCAATATATTGAAGAACAA 1066
QY 158 GluAlaThrValIleThrAspValProTyrAspIleValProGly-----Ala 173
DB 1067 ATGAAACATACCTGATATTAAACAAAGCATTTTTCAAATTCAAAGCGCAACAGCAAAATTC 1126
QY 174 LeuCyValLeuGlnGlnGlyGlyGluThrGlnIleLeuArgValGlyGluAlaLeu 193
DB 1127 ATTCGATCTACTACGTAGAGCAACAACTGAATCTTGAAGAGGCTCCACAGATTGAA 1186
QY 194 LeuIleGlu-----ArgLeuProIleSerPheIleAspValVal 207
DB 1187 CTTAAGGAATCAGAGGAATTCAGTCACTATTTAAACTTTTCAAGAACTGATGTC 1246
QY 208 LeuValIleAlaLeuAlaPheLeuAspValSerArgAspAlaMetAlaLeuAspProPro 227
DB 1247 GCGTATATGCTGCTGATGCTG-----CCCAAA 1273
QY 228 AspPheIleThrGlyTyrGluPheValGluGluAlaLeuIleLeuGlnGlnGly 247
DB 1274 GGACTTAATCTAATAT-----TATCGGATATGTTGAGTTAGCAAAAGCAAGCA 1327
QY 248 AlaSerSerLeuAlaProAspLeuArgAlaGlnIleAspGluThrLeu-----GluGluIle 266
DB 1328 ATTTGACCATTTTAGTAGTCTGCTGTCATCACTTGAGAGAGTTCTTATTAGTAATGTC 1387
QY 267 ThrProArgTyrVal----- 271
DB 1388 AAACCTACAGTATTTAAACCAATATATAGTAATATCACAACTTTTAATTACAAAGTA 1447
QY 272 -----LeuGlu 273
DB 1448 ACCAATGATATTAAAGATTGAAGCGGCACTAAGTCAAGCAATATTATTAATGATTGAA 1507
QY 274 LeuLeuGlyLeuProLeuGlyIleAspAspTyrAlaIleAspArgLeuAsn----- 289
DB 1508 TGATATTATGTTTCATTGGCGAGTGAAGGCTTTTGCAAAACATATATCAAAATTTTAT 1567
```

```

QY 290 -----GlyLeuSerGlyValArgAsnIleLeuTrpSerValGlyGlyGly----- 305
DB 1568 AAGGTGAATATTCCTCCCAATTAA-----GTAGTTAATCTGTGGTCAGAGATTCCACT 1624
QY 306 -----AlaSerAlaLeuValGlyGlyLeuThrArgGluIlePheMetAsnGlu 321
DB 1625 GTAGCAGGAATTCCTTGAGCTCATTCATCAACAAACGATGAGGTTATTAATAAAA 1684
QY 322 -----AlaPheLeuArgMetThrAlaAlaGluGlnValAspLeuPheValAlaThrPro 339
DB 1685 GCAATATGCAATTCGGAATGCTAAATGCAATGCAACAA-----CAAACA 1726
QY 340 SerAsnIleProAlaGluSerPheGluValTyrGluValAlaLeuAlaLeuAlaGln 359
DB 1727 GGTCAATTTAATACAGTAAATTTAGCAAAATTTTCAACAAATAGAAAGTTATAGAGCTG 1786
QY 360 AlaPheIleGlyIleIleAspProHisLeuLeuGlnAspAlaAspIleGlnGlnLeu 379
DB 1787 TAATTTATGCAAAA----- 1801
QY 380 GlnGlnAlaIleValIleMetAlaMetGlu-----IleProAla 391
DB 1802 TCACAAACAAAGAGTGCATCAATTTAGAAATTAAGTATCAAGAGGATATTATTCAGCT 1861
QY 392 MetLeuTyrAspThrArgAsnAsnTyrGluIleAspPheGlyLeuGluArgGlyLeuCy 411
DB 1862 TTAGCATTTGAT-----CAACGTGGTGCACTTA 1888
QY 412 AlaLeuLeuIleGlyIleVal-----ValAspGluCyArgMetTyr 424
DB 1889 AAAAGATGATGTCAGAACTCAATCTGAAACCAACAGTTGAACAAATAGAACATTA 1948
QY 425 LeuGlyLeuAspSerGluAsp-----SerGlnTyrArgAsnProAlaIleVal-----Glu 441
DB 1949 AAGTCTCTGTTCTGAGAAATTAATCAATATGCGCTTCAATTTATTAATGATCCAGAA 2008
QY 442 PheValLeuGluAsnSerAsnArgAspAspAsp----- 453
DB 2009 TATGTTTACACAGCATCAGATGCTCGAATTAAGTACGCGGACTATTAATCTTCATACGAA 2068
QY 454 -----AspLeuProGlyLeuCyIleLeu-----LeuGluThrTyrPhe 466
DB 2069 AAACGTGATATGATGATGATGCGAAAGGCTGTTGCCAGATTGCTTGGAGATGCTC 2128
QY 467 AlaGlyValValPheProArgPheArgAspThrIleAspIleValPheIleGlyAsp 486
DB 2129 GCG-----AAACGTTGAAGAGAGCAAGGCGCCAAATGCACTTAATTTTAACTT 2176
QY 487 TyrTyrAspAspProMetValLeuSerTyrLeuGluArgValGluValValGlnGlySer 506
DB 2177 TATTATGAT-----GTAGATGACACAGAAGAATT 2206
QY 507 ProLeuAlaAlaAlaThrMetAlaArgIleGlyAlaGluIleValIleValIleSerAla 526
DB 2207 AACATACAAAGAAACATATATTTGAAACGAATTTGTTGCAATGTTGCGCAAGATATT 2266
QY 527 MetGlnAlaLeuGlnIleVal-----PheProSerArgTyrThrAspArgAsnSerAlaGlu 545
DB 2267 CTTTCTCTTGGAAGTTTAAATATGACGACAAATATTCCTGACATTAAGTGCAGAA 2326
QY 546 -----ProIleAspValGlnGluThrVal----- 553
DB 2327 TTCGCTAAAGTTAAGCCAGTAAAGTTAATGAAAGATGTAATCTCTGAAGATCGT 2386
QY 554 PheSerValAsp-----ProValGlyAsnAsnValGlyArgAspGly 567
DB 2387 TTTAATGTGATGATCTTAAAGTTGAAGTACCTGTGAATATATGAAATTTTGTGAAAGATTT 2446
QY 568 GluProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsnAsp 587
DB 2447 TCAGAGGAGAAAGTTTATTAATCAAAAGAAAGCTGCACAACTATTCGCTGATCAAGAT 2506
```

QY 588 TyrAla-----IleArgAlaGlyValSerGluSerValAsp 600
 Db 2507 GCAGCTACTCATTACCATATTATTATTAGTGCAGGTGTATACAGAGATTGTCCAA 2566
 QY 601 GluThr 602
 Db 2567 GATACA 2572

RESULT 11

US-11-098-686-9088
 / Sequence 9088, Application US/11098686
 / Publication No. US20060024696A1
 / GENERAL INFORMATION:
 / APPLICANT: Kanur, Vivek and Gebhart, Connie J.
 / TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
 / FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING
 / FILE REFERENCE: 09531-128001
 / CURRENT APPLICATION NUMBER: US/11/098,686
 / PRIOR FILING DATE: 2005-04-04
 / PRIOR APPLICATION NUMBER: PCT/US03/31318
 / PRIOR FILING DATE: 2003-10-01
 / PRIOR APPLICATION NUMBER: US 60/416,395
 / NUMBER OF SEQ ID NOS: 11433
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 9088
 / LENGTH: 1908
 / TYPE: DNA
 / ORGANISM: Lawsonia intracellularis
 / US-11-098-686-9088

Alignment Scores:

Pred. No.:	0.14	Length:	1908
Score:	112.00	Matches:	124
Percent Similarity:	34.3%	Conservative:	86
Best Local Similarity:	20.2%	Mismatches:	213
Query Match:	2.8%	Indels:	190
DB:	12	Gaps:	32

US-10-600-070b-2 (1-801) x US-11-098-686-9088 (1-1908)

QY 235 PheValGluGluValLeuLysLeuLeuGluGluGluValAspSerLeuAlaProAsp 254
 Db 406 TTGGTGGCTCAGTCGACGAGAAATGCTTAACACAGGAAAGACGACATTGAAGAT 465
 QY 255 LeuArgAlaGlnIleAspGluThrLeuGluGluThr----- 267
 Db 466 GTT---GCAGGTGTGATGAAGCTAAAGAACTCTGAAAGTTGAGATTTTATCT 522
 QY 268 -----ProArgTyrValLeuGluLeuLeu 275
 Db 523 AATCCAGAAATTATACAGATTAGTGTGCTGATCTTAAGAGATCTT---CTTGTT 579
 QY 276 GlyLeuPro---LeuGlyAspAspTyrAlaAlaLysArgLeuAsnGlyLeuSerGlyVal 294
 Db 580 GGCCTCTCTGTCAGGGAACATTCCTGCTGACCTGTAGCTGTGAAGAGGGGTT 639
 QY 295 ArgAsnIleLeuTyrSerValGlyGlyGlyValAspSerAlaLeuValGlyGlyLeuThr 314
 Db 640 CCT-----TTTTCCTCATTTCTGCTGAT----- 666
 QY 315 ArgGluLysPheMetAsnGluAlaPheLeuArgMetThrAlaAlaGluGluValAspLeu 334
 Db 667 -----TTTGTGAGATGTTGTGTGTGTGAGAGCATTAAGTACGAGATCTC 714
 QY 335 PheValAlaThrProSerAsnIleProAlaGluSerPheGluValTyrGluVal---Ala 353
 Db 715 TTTTATACAGGTAAATAAATGCTCCATGCTTATTTT---ATTGATGAATTTGATGCA 771
 QY 354 LeuAlaLeuValAlaGlnAlaPheIleGlyLysLysProHisLeuLeuGluAspAlaAsp 373
 Db 772 GTAGAGCTGACGCTGCGCTGCGCTTGGC-----GGAGACATGAT 813

QY 374 LysGlnPheGlnGluGluGlnAlaLysValMetAlaMetGluIleProAlaMetLeu 393
 Db 814 GAAGAGAAACAAACCTTATATCACTTTAGTT----- 846
 QY 394 TyrAspThrArgAsnAsnTrpGluIleAspPheGlyLeuGluArgGlyLeuCyAlaLeu 413
 Db 847 -----GAAATGGAT---GGTTTGAAATCTTAACGAAGGTGTTAT 882
 QY 414 LeuIleGlyLysValAspGluCysArgMetTrpLeuGlyLeuAspSerGluAspSerGln 433
 Db 883 TTTATCGCAGCTACAAAT----- 900
 QY 434 TyrArgAsnProAlaIleValGluPheValLeuGluAsnSerAsnArgAspAsnAsp 453
 Db 901 -----CGTCCAGATGTTCTTGTCCAGCTCTATTAAAGCCTGAAGGTTTGTATACAA 954
 QY 454 -----AspLeuProGlyLeuCysLysLeuLeuGluGluThrTrpLeuAla 467
 Db 955 GTTGTGTTCTTACTCTGATGTAAAGACGTTTAAATTTCTTGA----- 1002
 QY 468 GlyValAlaPheProArgPheArgAspThrLysAspLysLysPheLysLeuGlyAspTyr 487
 Db 1003 -----GTCCATATCTGT---CGACACACCTGATTAAGATGTTATCTT----- 1044
 QY 488 TyrAspAspProMetValLeuSerTyrLeuGluArgValGluValVal---GlnGlySer 506
 Db 1045 -----GAAATTATAGCTACAGCACCA 1065
 QY 507 ProLeuAlaAlaAlaThrMetAlaArgIleGlyValGluHisValLysAlaSerAla 526
 Db 1066 CCAGGTTTTTCAGAGACGACGACTTGAAACCTT-----GTTAATGAAGCACCA 1113
 QY 527 MetGlnAlaLeuGlnLysValPheProSerArgTyrThrAspArgAsnSerAlaGluPro 546
 Db 1114 CTTCAGGCGACTTGA-----GATGATAAATTTAATTGATGATG 1152
 QY 547 LysAspValGlnGluThrValPheSerValAspProValGlyAsnAsnValGlyArgAsp 566
 Db 1153 AAAGATTTTGAG-----TACGCTAAAGACAAAGTT-----CTAATGGCTAAAGAG 1197
 QY 567 GlyGluProGlyValPheIleAlaGluAlaValArgProSerGluAsnPheGluThrAsn 586
 Db 1198 ---CGAGTAGTCTCATTTCTTCTGATGAGAAAGAAATTAACACCTTATCTGTAAGCT 1254
 QY 587 AspTyrAlaIleLeuArgAlaGlyLysSerGluSerValAspGluThrThrValGluMet 606
 Db 1255 GGCATGCACTT----- 1266
 QY 607 SerValAlaAspMetLeuLysGluAla-----SerValLysIleLeuAlaAla 622
 Db 1267 ---GTTGACAGTTTACTTCCAAAGACAGATCCAGTCCACAAAGTAACTATTCTCGA 1323
 QY 623 GlyValAlaIleGlyLeuIle-----SerLeuPheSerGln 634
 Db 1324 GGTGTGCTCTTGTGTGATGACATGATCAATTACTGAAGACATGACATGACTTATTTCAAG 1383
 QY 635 LysTyrPheLeuLysSerSerSerPheGlnArgLysAspMetValSerSerMetGlu 654
 Db 1384 GAATATTATTATTAATAACCTGATGCTGTTATTAGCTGTGACACAGCTGAAGAAATTATA 1443
 QY 655 SerAspValAlaThrIleGlySerValArgAlaAspAspSerGluAlaLeuProArgMet 674
 Db 1444 TTTGATACATATTACAACAGGTGA-----GTTAATGATTTGAACCTGCACCAATATG 1497
 QY 675 AspAlaArgThrAlaGluAsnIleValSerLysTyr-----GlnLysIleLysSer 691
 Db 1498 -----CCAGGAAATGTTGTGAAATGGGATGAGCGGCTTATAGCA 1545
 QY 692 LeuAlaPheGlyProAspHisArgIleGluMetLeuProGluValLeuAspGlyArgMet 711
 Db 1546 TTTATCAATTGCT-----GACCGTGAGAAAGGTTTTTATTTGGTAGAGAG 1590
 QY 712 LeuLysIleThrThrAspArgAlaAlaGluThrAlaGlnLeuGlyLeuValTyrAspTyr 731


```

QY 432 SerGlnTyrArgAsnProAlaIleValGluPheVal-----LeuGluAsnSer 447
    |||||
DB 6511 AGATGTACCGAACGGGTGATTGGTACCGTGTACCAATGGGAATTGTAATTATTA 6570
QY 448 AsnArgAspAspAsnAsp-----AspLeuPro 456
    |||||
DB 6571 GGAAGATATAGATACCAAGTCACAGATAGAGCGCTCCGAATTGAGCTTGGTAATTGAA 6630
QY 457 GlyLeuCyLeuLeuGluThrTrpLeuAlaGlyValValPheProArgPheArgAsp 476
    |||||
DB 6631 GCGCAATTGCTCATTTAGAAAGTGTAGTCACCAATTTGT--ACACAAATGACCGAT 6687
QY 477 ThrLeuAspLysLys-----PheLeuLeuLysAspTyrTyrAspAspPromet 492
    |||||
DB 6688 CATTAAGGTGCGCAACAAAGATTGCGCATATGTTGTGCGCAAAAGTATGATGA-- 6741
QY 493 ValLeuSerTyrLeuGluArgValGluValValGlnGlySerProLeuAlaAlaAla 512
    |||||
DB 6742 -----GACGAAATTCGT----- 6753
QY 513 ThrMetAlaArgIleGlyAlaGluHisValLysAlaSerAlaMetGlnAlaLeu---- 530
    |||||
DB 6754 -----GAAATGTGAAAAAGCGTTGCGCAATATATGTTACCA 6792
QY 531 -----GlnLysValPhePro-----SerArgTyrThrAspArgAsn 542
    |||||
DB 6793 AGTTATTTGCTTTCATGAAAGCGATTCCTTAATAATAAGAAAAAGTTGATGCGAAA 6852
QY 543 SerAlaGluProLysAspValGlnGluThrValPheSerValAspProValGlyAsnAsn 562
    |||||
DB 6853 CATTCGCATTCGGTTGATCTTTATGAAACAGATATGATACAGTCATGTGGGACACAGA 6912
QY 563 ValGlyArgAspGlyGluProGlyValPheIleAlaGluAlaValArgProSerGluAsn 582
    |||||
DB 6913 AACGAAAAAGAACCAATGCTTTCGCTT--ATTGGCAGAGCTTTGGGATAGAGAAAT 6969
QY 583 PheGluThrAspAspTyrAlaIleArgAlaGly----- 593
    |||||
DB 6970 ATCAGTGTTCACATTAATTTCTTAACTGTGTGTCATTCATTAATGCGACCAATTCG 7029
QY 594 ValSerGlu-----SerSerValAspGluThr 602
    |||||
DB 7030 GTATCAAAATTTATAGTGTTCGCCAGTGAAGATGCTCTTAATAATGCTTTCAGTAT 7089
QY 603 ThrValGluMetSerValAlaAspMetLeuLysGluAlaSerValLysIleLeu--Ala 621
    |||||
DB 7090 ACAACGTAGCTACATGACGCGGTGTAGAAAGATTTGGTAAAGCGCTGTGACGAA 7149
QY 622 AlaGlyValAlaIleGlyLeuIleSerLeuPheSerGlnLysTyrPheLeuLysSerSer 641
    |||||
DB 7150 GTAGCTGTAAACGAGCGCATTCACAAAGATCTACCGAGAACATTAATTCGATTTGTCG 7209
QY 642 SerSerPheGlnArgLysAspMetValSerSerMetGluSerAspValAlaThrIleGly 661
    |||||
DB 7210 TATTCACACAAAGAAATTTATTTCTTATCTACATGAGAAAGAAACCAATTAATTAAT 7269
QY 662 SerValArgAlaAspAspSerGluAlaLeuProArgMetAspAlaArgThrAlaGluAsn 681
    |||||
DB 7270 ATCTCTGTGCTGTGGAT-----ATTATGGAACACTAGAT----- 7305
QY 682 IleValSerLysTyrGlnLysIleLysSerLeuAlaPheGlyProAspHisAlaGlyLeu 701
    |||||
DB 7306 ---GTTACGCTATTGAAAAAGCAATCCAACTTAATAGAAAGAACCACTTCAATACGT 7362
QY 702 MetLeuProGluValLeuAspGlyArgMetLeuLysIleTyrThrAspArgAlaAlaGlu 721
    |||||
DB 7363 GCAACATTTGAATTCGTGACGCGCAACCTGTGCAATC----- 7401
QY 722 ThrAlaGlnLeuGlyLeuValTyrAspTyrThrLeuLeuLysSerValAspSerVal 741
    |||||
DB 7402 -----ATCCACGATGATATGAAATTCCTGTGCAATTTATTGACCTT 7443

```

```

QY 742 ThrValSerAlaAspGlyThrArg-----AlaLeuValGluAlaThrLeuGluGluSer 759
    |||||
DB 7444 ACTGTATGCCAGACAGATTCACGATAGAGAAAGTACATGATTAATGTTAAAAAGCTT 7503
QY 760 AlaCyLeuSerAspLeuValHis---ProGluAsnAsnAlaThrAspValArg 776
    |||||
DB 7504 AAAAGATATACATCATCTCCGAATATGTCCTGTATGATGATGTCACAAATTTGTAAG 7557
RESULT 14
US-10-513-786-8
; Sequence 8, Application US/10513786
; Publication No. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio
; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method for
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513,786
; PRIOR FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 11115
; TYPE: RNA
; ORGANISM: Bacillus cereus
US-10-513-786-8

Alignment Scores:
Pred. No.: 1.4 Length: 11115
Score: 111.50 Matches: 159
Percent Similarity: 33.7% Conservative: 117
Best Local Similarity: 19.4% Mismatches: 301
Query Match: 2.7% Indels: 241
DB: 8 Gaps: 36

US-10-600-070b-2 (1-801) x US-10-513-786-8 (1-11115)
QY 105 IleArgArgAlaPheGluAlaArgValSerLysProProGlnPheGlyPheSerAspAsp 124
    |||||
DB 5389 AUTCAGAAUUGUUGCAGAGAGUGUUCACAAAACCGCA-----UCUGCAAC 5436
QY 125 AlaLeuIleSerArg--ArgGlnIleLeuGlnAlaAlaCySerGluThrLeuSerAsnPro 143
    |||||
DB 5437 GCGCUUATUUUUGCAACAAAGAGUAGACAUACGAGAGUUAUUGCGCAGCAAAUUGCA 5496
QY 144 ArgSerArgArgGluTyrArgGlnGlyLeuLeuAspAspGluGluAlaThrValIleThr 163
    |||||
DB 5497 CUUGCAUAAUUAAGUUGCAGAGUGUUGACCAAAACCGAGUAGUUGCGUGUUAUGU 5556
QY 164 AspValProTyrAspLysValProGlyAlaLeuCyValLeuGlnGlyGly----- 181
    |||||
DB 5557 GAUCGUCUUCUUGCAAGUGUGUGUUAUUAUGGUCGUUUUUAAGCAGUGUGUGGUU 5616
QY 182 -----GluThrGlnIleValLeuArgValGlyGlu 191
    |||||
DB 5617 UUGCAUUGUAGUACAGCGUACCCGAGUACAGAACGAAUACGUCUGCAAAUUAUGAG 5676
QY 192 Ala-----LeuLeuLysGlnArgLeuProLysSerPheLys 203
    |||||
DB 5677 GCAACUUAUUCUUAACAAAGAGUAGUUAUUCUUAAGAGUUCUUA-----GAUUUGAG 5730
QY 204 GlnAspVal---ValLeuValMetAlaLeuAlaPheLeuAspValSerArgAspAlaMet 222
    |||||
DB 5731 GGGGAAGUUUUUUAUCUUAAGUAGUAGACAGUCUUUGAAGGGAGUAGAAAGUUUAACA 5790
QY 223 AlaLeu-----Asp 225
    |||||
DB 5791 AAUAUUAUAUUAUUCUUAACUUGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5850

```


D_b 1013 GGTGGCTGAGATGAGCTTCGGCTGCGAGACACCTGGCTTAAGCTTCTTCAATGAGCAG 1077

Db 1871 GAGTCCGATTCGAGACCCCTGGTGAGGACGACGCCA---GAGGAAGAGTACTACTCTTACC 1927
:::|::|
:::|::|
:::|::|

Cv 516 AGCCTTGGTIV-----AAAGGHHI[ELVA]LJWVLA 524
- - - - -

```
Db      1928 AAACCTGGCAACACCATTTCCAGCCTGTTTGGCGGTTGATACCTCATCAGATGCCAAAGAG 1987
      :::::||||
Qy      525 SerAlaMetGlnAlaLeuGln-----LysValPheProSerArgTyrThrAspArgAsn 542
      :::::|||||
Db      1988 AATGCTACTGATGCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2047
      AATGCTACTGATGCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      543 SerAlaGluProLysAspValGlnGlnIuThrValPheSer----- 555
      |||||
Db      2048 CCTGCAGAACAGGGGGAGACTCAAGAGAGAACTGAAGCCCAATGAGAGATACCTCCAG 2107
      |||||
Qy      556 -----ValAspProValGlyAsnAsnValGlyArgAspGlyGlu----- 568
      :::::|||||
Db      2108 CCTCCACCTCTGAGGCTTAAGGG---GATGCAGCCCGTGAAGGAGAGAAACACCTGATGAA 2164
      :::::|||||
Qy      569 ---ProGlyValPheIleAlaGlnAlaValArgProSerGluAsnPheGlnIuThrAsnAsp 587
      :::::|||||
Db      2165 AAAGAAAGTGGGAGCAAGTCTGAGGCCCAATGAGAAAGGCGCAGCGGGGCT 2224
      :::::|||||
Qy      588 TyrAlaIleArgAlaGlyValSerGluSerSer----- 598
      :::::|||||
Db      2225 GAGGCTGTCCTCCAGCTCCCGAGAGAAAAAAGCAGAACTGCGCGGAAACAGAAA 2284
      :::::|||||
Qy      599 ---ValAspGluThrThrValGlnMetSerValAlaAspMet 611
      |||||
Db      2285 ATGCTGAGAGAGATAGTGTGAACTGCTGCTTGTGACCTG 2326
      |||||
```

Search completed: February 21, 2006, 14:51:15
Job time : 3423 secs

gene
/db_xref="taxon:3702"
/clone="GSRLS892C08"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. .919
/gene="At5g42480"

ORIGIN

Query Match 37.5%; Score 902.2; DB 4; Length 919;
Best Local Similarity 99.7%; Pred. No. 4,7e-244;
Matches 904; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATGGAAGCTCTGATCGTCAAGTGTCTCTCCCATTTCAATATGCGATTACCA 60
DB 13 ATGGAAGCTCTGATCGTCAAGTGTCTCTCCCATTTCAATATGCGATTACCA 72
QY 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACTCTACAACTATCTGCTCCGACG 120
DB 73 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACTCTACAACTATCTGCTCCGACG 132
QY 121 AAATGGGCGACGCTCTCTCTCCGACTTCAATTTCACTTCGATTCTCTCTCTCTCC 180
DB 133 AAATGGGCGACGCTCTCTCTCCGACTTCAATTTCACTTCGATTCTCTCTCTCTCC 192
QY 181 TTGGCCACCGCCACACACCGCCACTCTGCTCTCTGACCATCTATTGANTCGTCC 240
DB 193 TTGGCCACCGCCACACACCGCCACTCTGCTCTCTGACCATCTATTGANTCGTCC 252
QY 241 GAAAGCCACGCTCCCATCCCATTTGATTTTCAACAGTATTTAGAGCTCAACACATTTTC 300
DB 253 GAAAGCCACGCTCCCATCCCATTTGATTTTCAACAGTATTTAGAGCTCAACACATTTTC 312
QY 301 TTAACCGATGGAATCAGAAAGCATTTGAAAGCTAGGTTTGAACCGCGCAATTCGCT 360
DB 313 TTAACCGATGGAATCAGAAAGCATTTGAAAGCTAGGTTTGAACCGCGCAATTCGCT 372
QY 361 TTACAGCAGACGCTTTAATCAAGCGGAGACAGATTTCTCAAGTGTCTTGGAAACCTG 420
DB 373 TTACAGCAGACGCTTTAATCAAGCGGAGACAGATTTCTCAAGTGTCTTGGAAACCTG 432
QY 421 TCTAATCTCGTCTAAGAGAGATACATGAAAGTCTTCTGATGATGAAGAGCTACA 480
DB 433 TCTAATCTCGTCTAAGAGAGATACATGAAAGTCTTCTGATGATGAAGAGCTACA 492
QY 481 GTCAATCACTGATGTTCTTGGGATAGGTTCTGGGCTCTCTGATGATGCAAGAGT 540
DB 493 GTCAATCACTGATGTTCTTGGGATAGGTTCTGGGCTCTCTGATGATGCAAGAGT 552
QY 541 GGTGAGCTGAGATAGTTCTTCGGGTTGGAGGCTCTGCTTAAGAGAGGTTGCCCTAG 600
DB 553 GGTGAGCTGAGATAGTTCTTCGGGTTGGAGGCTCTGCTTAAGAGAGGTTGCCCTAG 612
QY 601 TCGTTAAGCAAGATGTTGTTTATGATAGGCGTTGCTGATGCTCGAGGAGAT 660
DB 613 TCGTTAAGCAAGATGTTGTTTATGATAGGCGTTGCTGATGCTCGAGGAGAT 672
QY 661 GCTATGCAATGATCACTGATTTTATTACTGTTATGATGATTTGTTGAGAAAGCTTTG 720
DB 673 GCTATGCAATGATCACTGATTTTATTACTGTTATGATGATTTGTTGAGAAAGCTTTG 732
QY 721 AAGTTTTTACAGAGAGAGCAAGTACCTTCAACCGATTTTACCTGCAACAAATGAT 780
DB 733 AAGTTTTTACAGAGAGAGCAAGTACCTTCAACCGATTTTACCTGCAACAAATGAT 792
QY 781 GAGACTTTGGAAGAGATCACTCCGCTTATGTTTGAAGCTACTTGCTTACCGCTTGT 840
DB 793 GAGACTTTGGAAGAGATCACTCCGCTTATGTTTGAAGCTACTTGCTTACCGCTTGT 852
QY 841 GATGATTAACGCTGGAAGAAAGCTAATGTTTAAAGCGTGTGCGAAATATTTGTGCTCT 900
DB 853 GATGATTAACGCTGGAAGAAAGCTAATGTTTAAAGCGTGTGCGAAATATTTGTGCTCT 912

```

QY 901 GTTGGAG 907
DB 913 GTTGGAG 919

RESULT 2
CNS09YNM
LOCUS
DEFINITION

CNS09YNM 741 bp mRNA linear HTC 06-FEB-2004
GSRLS11592A05 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION BX833489.1 GI:42455179

VERSION HTC; GSRLT cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 741)
Castelli V., Aury J.M., Jailion O., Wincker P., Scarpelli C., Schachter V.,
Menard M., Cruaud C., Quetier F., and Salanoubat M.,
Temple G., Caboche M., Weissenbach J. and Salanoubat M.,
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

AUTHORS Unpublished
TITLE 2 (bases 1 to 741)

JOURNAL Direct Submission
REFERENCE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES Location/Qualifiers
1. .741
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSRLS11592A05"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. .741
/gene="At5g42480"

ORIGIN
Query Match 26.3%; Score 633; DB 4; Length 741;
Best Local Similarity 100.0%; Pred. No. 1.1e-167;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1774 GCTGGGCTCTCAAGAGTAGCGTTGTAAGAACTGTTGAATGTCCTGTTGCTGATG 1833
DB 1 GCTGGGCTCTCAAGAGTAGCGTTGTAAGAACTGTTGAATGTCCTGTTGCTGATG 60

QY 1834 TTAAGAGGCAAGTGAAGATCTAGCTGCTGTTGGCAATTTGCAATTTCACTG 1893
DB 61 TTAAGAGGCAAGTGAAGATCTAGCTGCTGTTGGCAATTTGCAATTTCACTG 120

QY 1894 TTACAGCAGAGATATTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGTTCT 1953
DB TTACAGCAGAGATATTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGTTCT 1953

Db 121 TTCAGCCAGAGTATTTCTTAAAGAGCTCATCTTTCAAGCAGATATATGTTCT 180
 Qy 1954 TCTATGGAATCTGATGCTGACATAGGCTGCTGAGCTGACGATTCAGAAAGCACT 2013
 Db 181 TCTATGGAATCTGATGCTGACATAGGCTGCTGAGCTGACGATTCAGAAAGCACT 240
 Qy 2014 CCCAGATGATCTGAGCTGAGAGATATAGTATCCAAAGTGCAGAAATTAAGTCT 2073
 Db 241 CCCAGATGATCTGAGCTGAGAGATATAGTATCCAAAGTGCAGAAATTAAGTCT 300
 Qy 2074 CTGGCTTTTGGGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGGCGAATG 2133
 Db 301 CTGGCTTTTGGGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGGCGAATG 360
 Qy 2134 CTGAATTTTGGATCTGACAGAGAGCTGAAATCTCCGAGCTTGGTGTATGATAT 2193
 Db 361 CTGAATTTTGGATCTGACAGAGAGCTGAAATCTCCGAGCTTGGTGTATGATAT 420
 Qy 2194 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTG 2253
 Db 421 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTG 480
 Qy 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAAT 2313
 Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAAT 540
 Qy 2314 GCTACTGATGTCGAACCTACACAAAGATAGAAATTTTCTGCTCAAGTCAAGGTGG 2373
 Db 541 GCTACTGATGTCGAACCTACACAAAGATAGAAATTTTCTGCTCAAGTCAAGGTGG 600
 Qy 2374 AAAATCACTGAAGGCTCTGTTCTTGATCATTA 2406
 Db 601 AAAATCACTGAAGGCTCTGTTCTTGATCATTA 633

RESULT 3
 CENS09255 741 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSJ751L12E09 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).
 BX833051.1 GI:42455503
 HCT; GSJ751L12E09
 ACCESSION Arabidopsis thaliana (thale cress)
 VERSION Arabidopsis thaliana
 KEYWORDS Arabidopsis thaliana
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 741)
 Castelli V., Aury J.M., Jailion O., Wincker P., Clepet C.,
 Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V.,
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 741)
 Genoscope.
 Direct Submission
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 location/Qualifiers
 FEATURES
 source
 1..741
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSJ751L12E09"
 /cissue_type="Silique"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1..741
 /gene="At5g42480"

ORIGIN
 Query Match 26.3%; Score 633; DB 4; Length 741;
 Best Local Similarity 100.0%; Pred. No. 1.1e-167;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1774 GCTGGGGTCTCAGAGAGTACGCTGATGAAACTGTTGAATGCCGTTGCTGATATG 1833
 Db 1 GCTGGGGTCTCAGAGAGTACGCTGATGAAACTGTTGAATGCCGTTGCTGATATG 60
 Qy 1834 TTAAAGAGGCAAGTGTGAAGATCTAGCTGCTGTGTGGCAATTGAGATTTTCACTG 1893
 Db 61 TTAAAGAGGCAAGTGTGAAGATCTAGCTGCTGTGTGGCAATTGAGATTTTCACTG 120
 Qy 1894 TTGAGCAGAGATTTTCTTAAAGACGCTCATCTTTTCAACGCAAGATATGTTTCT 1953
 Db 121 TTCAGCAGAGATTTTCTTAAAGACGCTCATCTTTTCAACGCAAGATATGTTTCT 180
 Qy 1954 TCTATGGAATCTGATGCTGACATAGGCTGACAGCTGACAGATTCAGAAAGCACT 2013
 Db 181 TCTATGGAATCTGATGCTGACATAGGCTGACAGCTGACAGATTCAGAAAGCACT 240
 Qy 2014 CCCAGATGATCTGAGCTGAGAGATATAGTATCCAAAGTGCAGAAATTAAGTCT 2073
 Db 241 CCCAGATGATCTGAGCTGAGAGATATAGTATCCAAAGTGCAGAAATTAAGTCT 300
 Qy 2074 CTGGCTTTTGGGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGGCGAATG 2133
 Db 301 CTGGCTTTTGGGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGGCGAATG 360
 Qy 2134 CTGAATTTTGGATCTGACAGAGAGCTGAAATCTCCGAGCTTGGTGTATGATAT 2193
 Db 361 CTGAATTTTGGATCTGACAGAGAGCTGAAATCTCCGAGCTTGGTGTATGATAT 420
 Qy 2194 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTG 2253
 Db 421 ACACTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTG 480
 Qy 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAAT 2313
 Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAAT 540
 Qy 2314 GCTACTGATGTCGAACCTACACAAAGATAGAAATTTTCTGCTCAAGTCAAGGTGG 2373
 Db 541 GCTACTGATGTCGAACCTACACAAAGATAGAAATTTTCTGCTCAAGTCAAGGTGG 600
 Qy 2374 AAAATCACTGAAGGCTCTGTTCTTGATCATTA 2406
 Db 601 AAAATCACTGAAGGCTCTGTTCTTGATCATTA 633

RESULT 4
 A1998415/c 561 bp mRNA linear EST 08-SEP-1999
 LOCUS 701545606 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
 DEFINITION thaliana cDNA clone 701545606, mRNA sequence.
 ACCESSION A1998415
 VERSION A1998415.1 GI:5845320
 KEYWORDS EST.

QY	2141	TTTGGACATGACAGACAGCTGAACTCGCGACCTTGGGTGGTTATGATTAACACGT	220
QY <td>2141</td> <td>TTTGGACATGACAGACAGCTGAACTCGCGACCTTGGGTGGTTATGATTAACACGT</td> <td>220</td>	2141	TTTGGACATGACAGACAGCTGAACTCGCGACCTTGGGTGGTTATGATTAACACGT	220
Db <td>242</td> <td>TTTGGACATGACAGACAGCTGAACTCGCGACCTTGGGTGGTTATGATTAACACGT</td> <td>301</td>	242	TTTGGACATGACAGACAGCTGAACTCGCGACCTTGGGTGGTTATGATTAACACGT	301
QY <td>2201</td> <td>TGAAACTATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCCGCTCTGTGTGAAG</td> <td>2260</td>	2201	TGAAACTATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCCGCTCTGTGTGAAG	2260
Db <td>302</td> <td>TGAAACTATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCCGCTCTGTGTGAAG</td> <td>361</td>	302	TGAAACTATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCCGCTCTGTGTGAAG	361
QY <td>2261</td> <td>CAACTCTGGAGAGTCTGCTTGTCTATCTGATTTGGTTTCATTCAGAAAAACAATGCTACTG</td> <td>2320</td>	2261	CAACTCTGGAGAGTCTGCTTGTCTATCTGATTTGGTTTCATTCAGAAAAACAATGCTACTG	2320
Db <td>362</td> <td>CAACTCTGGAGAGTCTGCTTGTCTATCTGATTTGGTTTCATTCAGAAAAACAATGCTACTG</td> <td>421</td>	362	CAACTCTGGAGAGTCTGCTTGTCTATCTGATTTGGTTTCATTCAGAAAAACAATGCTACTG	421
QY <td>2321</td> <td>ATGTAGAAACCTACACAACAATATGCAAGTTTCTGTGTCGAAGTCAGAGTGGAAAAATCA</td> <td>2380</td>	2321	ATGTAGAAACCTACACAACAATATGCAAGTTTCTGTGTCGAAGTCAGAGTGGAAAAATCA	2380
Db <td>422</td> <td>ATGTAGAAACCTACACAACAATATGCAAGTTTCTGTGTCGAAGTCAGAGTGGAAAAATCA</td> <td>481</td>	422	ATGTAGAAACCTACACAACAATATGCAAGTTTCTGTGTCGAAGTCAGAGTGGAAAAATCA	481
QY <td>2381</td> <td>CTGAAGGCTCTGTTCTTGTCATATTA</td> <td>2406</td>	2381	CTGAAGGCTCTGTTCTTGTCATATTA	2406
Db <td>482</td> <td>CTGAAGGCTCTGTTCTTGTCATATTA</td> <td>507</td>	482	CTGAAGGCTCTGTTCTTGTCATATTA	507
RESULT 6			
LOCUS	CL965374	2307 bp	DNA
DEFINITION	OsIFCC012338 <i>Oryza sativa</i> Express Library Oryza sativa (indica		
ACCESSION	CL965374		
VERSION	CL965374.1	GI:52385433	
KEYWORDS	GSS.		
SOURCE	<i>Oryza sativa</i> (indica cultivar-group)		
ORGANISM	<i>Oryza sativa</i> (indica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhacridae; Oryzaceae; <i>Oryza</i> . 1 (bases 1 to 2307)		
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.		
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80486676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.		
FEATURES			
source	Location/Qualifiers 1..2307 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences"		
ORIGIN			
Query Match	18.6%;	Score 447;	DB 10;
Best Local Similarity	53.7%;	Pred. No. 1e-114;	
Matches 1256;	Conservative 0;	Mismatches 880;	Indels 201; Gaps 8;
QY	108	CTGCTCCGCGACGAAATGGCGACCGCTTCTCTCGACATTCATTTACCTCCGATTC	167
Db	132	CTGCGCGCAAGCGCTTTCGCGCAATTCACCTCTCTCCACGCGCGCTTCGACCC	191
QY	168	CTCTCTCTCTCTCTTTCGCGACGCGACGACGACGCACTGCTGCTCTGACGACATC	227
Db	192	GCGGCTCCCGGCTT	251
QY	228	TATGATGCTCCGCAAGCGACGTCCTCCATCCCATGATTTCTACAGGAT-----	280

Db	252	CCCGAGCGCCGCCAAGCGCTCCCTCCGCTCCAAAGTCGATTTCTAACAAGGTGAGGCCCG	311
Qy	281	-----TAGAGCTCAAAACACATTTCTTAAACCGATGAAATCAAGAGAC	323
Db	312	TGAATCTGCTCGGTTCTTAGGGGGCAGAGCCACATTTCTTGCGCGATGCAATCAGAGAGGC	371
Qy	324	AATGGAACCTTAGGGTTTGGAAAACCGCGGCATTTGGGTTTCAGCGACGACGGTTTAATCAG	383
Db	372	GTTCCAGGCACGAATAGCCAAAGCCACCGCAGATATGGCTACACGACCGATGCTCTGTGGG	431
Qy	384	CCGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTGTCTAATCCTCGCTTAGAAGAGA	443
Db	432	TGCTGCAAAATGCTGACGATTTGGCCATGACACTCTCATTGAACCAAGACTCCCGCATCA	491
Qy	444	GTACATGAAAGTCTTCTTGTATGATGAAGAAGCTACAGTCACTGATGTTCTTGGGA	503
Db	492	GTATGATGTCGCTTCTTGAGAAACCGGAAGAAGCTCTCACATGATATATTCCTTGGGA	551
Qy	504	TAAAGTTCTTGGGCTCTCTGTGATTTGCAGAAAGTGTGAGACTGAGATATGTTCTTCG	563
Db	552	CAAGAGGCTGG-----GAGGACCTGCTGCTGTG	584
Qy	564	GATTGATGAGCTCTGCTTAAGAGAGGTTGCTTAAGTCTTAAACAAGATGTGTTT	623
Db	585	AACGTGAGAAACAGTTGCTCTGAGTCGGCCACCCAGGCGTTTCAAGCAGAGAGTGTGCT	644
Qy	624	AGTTATGCGCTTGCGTTCTCGATGTCGTGAGGAGTCTATGGCATTTGATCCACCTGA	683
Db	645	AGCAGTGCCTGGCTTATGTGATCTATCAAGGGATGCTATGGCAGCAAGCCCTCCAGA	704
Qy	684	TTTATTACTGTTATGATTTGTTGAGGAAGCTTTGAAGCTTTTACAGAGAGAGAGC	743
Db	705	TGTATATGGCTGCTCGCGAGTGTCTGAGAGGGCTCTCAAGCTCTTGCAGAGAAATGAGC	764
Qy	744	AAGTAGCTTGACCCGATTTACGTCACAAATTTGATAGACTTTGGAAGAGATCACTCC	803
Db	765	AAGCAATTCGCACTGATCTGCTTTCACAGATTTGATAACTCTCGAGAGATTAACCC	824
Qy	804	GCGTTATGCTTTGGAGCTACTTGGCTTACCGCTTGATGATTTACGCTCGAAGAAAGCT	863
Db	825	TGCGTGTATTTGAGACTTCTCCCTTCCATTTGACACAGAGCATATAAGAGCGCA	884
Qy	864	AAATGCTTTAAGCGGTGCGAAATATTTGTGTCTGTTGAGAGAGGTGAGCATCAGC	923
Db	885	AGAAGGCGTTCAAGGTGCGAAGAACATTTTGTGAGCGCTTGACAGAGAGGTATTTGTC	944
Qy	924	TCTGTGGGGGTTGACCCGTCGAGAAATTATGATGAGGCGTTTTTACAAATGACGC	983
Db	945	CGTTGAGAGAAATTTCTCGTGAAGCCTTATGAACAGGCTTTTTTGAAGATGACATC	1004
Qy	984	TGCTGAGCAGGTATCTTTTGTATGCTACCCCAAGCAATATTCAGCAGAGTCAATTGA	1043
Db	1005	AATTGAACAGATGATTTCTTTCAAAAACCGAATATGACTTCTCTGAATGTTTGA	1066
Qy	1044	AGTTTACGAATGCACTGCTCTTGTGGCTCAAGCTTTTATGTGTAAGAACCAACACT	1103
Db	1065	AATTTACAATGATACACTTGACATGTGCTCAAGCAATTTAATGATAAAGGCCACAAT	1122
Qy	1104	TTTACAGATGCTGTAAGCAATTCAGCAACTTCAGCAGGCTAAGATATGCGTATGGA	1166
Db	1125	CATCATATAGCGGATGATCTTTTGAACAATCCAGAAAGTTCAA-----	1165
Qy	1164	GATTCCTGCGAATGTTGATGATACAGGAATTAATTTGGAGATATGACTTCGCTAGAAAG	1222
Db	1170	-----CATAGTTCTCATTAATGCTATGATTAATGATATGAGATGAGACTTCGATGGAAAG	1220
Qy	1224	GGAAGCTGTGCACTGCTTAATAGCAAAAGTTGATGAATGCCGATATGTGTGGGCTTGA	1283
Db	1221	GGCAATTCGTGCAATGCTAGTCGAGATGTTAGCAAGTCAGATATGTGGCTTGAATTTGA	1280
Qy	1284	CAGTGAGATTTACAAATTAAGGAATCCAGCTATTTGTGAGATTTGTTTGGAGAAATTC---	1340
Db	1281	TAAATGAGCTTCAACCATACAGAGACCCCAAAATTTCTAGAGTTTATTTGACCAACTCTAG	1340

QY 1341 AAATCGTATGACAAATGATCTCCTGGACATATGCAAAATGTTGGAAAACCTGCTGGC 1400
 |||||
 DB 1341 CATCAGTGAAGAAATATCTCTTCCAGGGCTGTGCAAGCTTTTGGAGACTTGCTTAT 1400
 QY 1401 AGGGGTTCTCTTCTTCTAGTTCAGAGACACCAAAAGATTAATACTCGGGAGCTA 1460
 |||||
 DB 1401 CTTTGAGGTTTCTTCTAGAGACAGATTAATCTCGGGGATCAGTTGACACTTGGAGATTA 1460
 QY 1461 CTATGATGATCTATGTTTGGAGTTACTTGGAAAGGTGAGGTACTTCAAGGTTCTCC 1520
 |||||
 DB 1461 CTAGAGATGATCCAGAAATTTTAACTTACCTAGAAAGAGTGAAGGTGTGTTCTTCA 1520
 QY 1521 TTTAGCTGCTGCTGCACTATGCGAAGATTTGAGCCGAGCATGTGAAGCTAGTCTAT 1580
 |||||
 DB 1521 TTTGGCTGCTGCTGCTGCTATGCAAACTTGCTCAAGCTACAGCTGCACTGGTGA- 1579
 QY 1581 GCAGGCACTGCAAGAAATTTTCTTCCCTGCTATACAGATGAAAATCTGGCTGAACCA 1640
 |||||
 DB 1580 -----CTGTGAATTCAAA 1592
 QY 1641 GGAATGCAAGACAGCTTTTATGATGATCTGTTGTGAACAATGAGCCGTGATG 1700
 |||||
 DB 1593 TGCTATTCAAGCG----- 1605
 QY 1701 TGAGCTGTGTCTTTATTGCAAGAGCTGTAGACCTCTGAAAACCTTGAACCTAATGA 1760
 |||||
 DB 1606 -----TTCAACAAGGTTTTCCTTGTATGATAGAACAGTTAGACAGTCAAGCTCAGC 1649
 QY 1761 TTATGCAATTGAGCTGGGCTCTCAGAGAGTACGGTTGATGAAAATCTACTGTGAATGTC 1820
 |||||
 DB 1650 CATGAAAATATCTAAAGATGGCCCTGGGGGATATCTTGAAAATTTTGAACAGAAAATGC 1709
 QY 1821 CGTTGCTGATATGTTAAAGAGCAAGTGTGAATACTCTAGCTGCTGTGTGGAAATGG 1880
 |||||
 DB 1710 ACCGCTCATGATTGCAAAATGCGCCTTGAAATATATCTGTCTGCGCATGTTTGC 1769
 QY 1881 ACTGATTTCACTGTTCAAGCAGAAATATTTTCTTAAAGACGCTCATCTTTTCAACGCAA 1940
 |||||
 DB 1770 ACTGTTGGCAGTATTTGGGCGCAAAATTTTCCCTGTAAGAGGCCCTTTCTGTATTAAG 1829
 QY 1941 GGAATAGGTTTCTTCTATGAAATCTGAT-----GTGCTACCAATAGGGTCACTCAG 1991
 |||||
 DB 1830 GAGTGAACATGATCTGTGGCAGTTGCTAATAGTGTGATCTACTGATATCTCTGACT 1889
 QY 1992 AGCTGAGATTTCAAGACATCTCCCAATGATGATGATGATGATGATGATGATGATGATG 2051
 |||||
 DB 1890 AGATGAAGATCCAGATACATATCTTGAAGATGATGCAAGCTGGCAGAAAGATTTGTCG 1949
 QY 2052 CAAGTGGCAGAAATTAAGTCTCTGGCTTTTGGGCTGATGACCGCATAGAAATGTTACC 2111
 |||||
 DB 1950 CAAGTGGCAGAAATTAAGTCTCTGGGCTTGGGCAAGAACTTGGGCTTGGATATTTGCA 2009
 QY 2112 AGAGTTTGGATGGGCAATGCTGAAGATTTGAGTGAACAGACAGCTGAATCTGGCA 2171
 |||||
 DB 2010 AGAGTTCTTGATGGCAACATGCTAAGATGTGAGTGAACGAGCAGCGAGATTTAGCG 2069
 QY 2172 GCTTGGGTTGGTTATATATTAATTAATCTGTTGAAATCTATCTGTTGACGTGTGACGCTC 2231
 |||||
 DB 2070 TCAATGGGTGTTCTGGAGATTAATCACTATCGATGATGACATGATGATGATGATGATG 2129
 QY 2232 AGCAGATGGAACCGGTCTGTTGAGGAAGCACTGGAAGAGTCTGTTGATCTATCTGA 2291
 |||||
 DB 2130 CTTAGATGATGAGAGGAGTGTGAGGCTTACATTTGATGATGAGGAGGAGGAGGAGGAGG 2189
 QY 2292 TTTGGTTTCATCCAGAAAACATGCTACTGATGTCAGAACCTTACCAACAAATGATGAGT 2351
 |||||
 DB 2190 TGTATTAGAGGCCAGAAACATGATTCATATGACAAAATATACACTACCGGATGATGAT 2249
 QY 2352 TTTCTGATCCAG---TCAGGGTGAAGAAATCACTGAAGGCTCTTCTTGTGATATA 2405
 |||||
 DB 2250 GGCCTTCTCCAGTGAAGAGGTGGAAGATTAACGAAAGAGACAGTCTCTCAAGTCTGA 2306

RESULT 7
 CO071968 813 bp mRNA linear EST 15-JUN-2004
 LOCUS GR_Ea30N03.r GR_Ea Gossypium raimondii cDNA clone GR_Ea30N03 3',
 DEFINITION mRNA sequence.
 ACCESSION CO071968
 VERSION CO071968
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 REFERENCE 1 (bases 1 to 813)
 Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
 Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and
 Wing, R. A.
 TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rwing@genome.arizona.edu
 Plate: 30 row: N column: 03.
 FEATURES
 source location/Qualifiers
 1..813
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Ea30N03"
 /ribose_type="whole seedlings"
 /dev_stage="first true leaves"
 /lab_host="PH10B"
 /clone_1lb="GR_Ea"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV. Library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Colonies
 plated/picked by AGI. More glycero clones held in -80."
 ORIGIN
 Query Match 18.1%; Score 436; DB 7; Length 813;
 Best Local Similarity 72.3%; Pred. No. 9,7e-112;
 Matches 584; Conservative 0; Mismatches 215; Indels 9; Gaps 1;
 QY 107 TCTGCTCCGCGCAGCAAAATGGGCGGACCGTCTTCTTCCGACTTCAATTTCACCTCG-- 163
 DB 5 TCTGCTCCGCGCAGTAAATGGGCGGAGGCGCTCTTGTGATCTTCAATTTCCTCCGCGC 64
 QY 164 -----ATTCTCTCTCTCTCTCTCTTGGCCACCGCCACCAACCGGCACCTCTGTCCTC 217
 DB 65 CCGATTAATTCCTGCTCTCTCTCTCTCTTCTTCAACCGCCACTCTTCTCCCTTACCTC 124
 QY 218 TGGCACAATCTATTGATTCGTCGCGAAGCGCAAGTCCCAATCCCATTTGATTTTTCACAG 277
 DB 125 CCTTCTCGCCCTTCTCTCTCCGGAAGCGCAGCTTTCATTCCTCTGATTTCTTCAAG 184
 QY 278 TATTAGAGCTCAAAACATTTCTTAACCGATGAAATCAAGAGATTCGAAGCTAGAGG 337
 DB 185 TTTTGAAGAGCGAGACTCATTTCTTAGTGATGATGATGATGATGATGATGATGATGATG 244
 QY 338 TTTGGAACCGCGCAATTCGGTTTCAAGCAAGACGCTTAACTACCGGAGCAATTC 397
 DB 245 TTTTGAACCGCGCTCAATATGGGTTCAATCAACACCAATTAATTCAGGAAGCAATTC 304
 QY 398 TTTCAAGCTGTTGCAAACTCTGCTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
 DB 305 TTTTACTGCTCTGTAAGACCTTATCTTAACCTTGGCTCTAGACGAATTTACACCAAGTC 364
 QY 458 TTCTTATGATGAAGAGCTACAGTCACTGATGATGATGATGATGATGATGATGATGATG 517

```

Db      365 TTGTCAGACGATGACGTGACACTATCATCACTCAAGTCCCTGGGACAAAGGTTCTCTGAG 424
Qy      518 CTCTCTGTATTTGCAAGAAAGGTGTGAGACTGAGATGTTCTTGGGTTGTGTAGGCTC 577
Db      425 CATGTGCGGTGTCGCAAGAAAGCTGGGAGACTGAAAGTGTGTCTTCAATTGGGAGAGTT 484
Qy      578 TGTCTAAGAGAGAGGTGCTTAAGTGTGTTTAAAGAAAGTGTGTGTTTAAAGGCGCTG 637
Db      485 TGTCTAAGAGAGAGGTGCTTCAAGGCAATTTAAACGAGATGTCTTGTGCAATGAGCCTTG 544
Qy      638 CGTTTCTGATGCTCTGAGGAGATCTATGGCATTTGATCCACTGATTTTATTAAGTGT 697
Db      545 CTATATGAGACTGTTCTGAGGAGTGTCTATGCTTGTGATTCCTCCAGATTTATGTTGTGTT 604
Qy      698 ATGAGTTTGTGAGGAAGCTTTGAAGCTTTTACGAGAGAAAGCAAGTACCTTGAC 757
Db      605 GTGAGTGTGTCGAGAGGCTTTGAAAGCTGTGACGAGAGAAAGTGTGCAAGCTTGTCTC 664
Qy      758 CGGATTTACGTGCAAAATTTGATGAGACTTTGAAAGAGATCACTCCGCTTATGCTTGG 817
Db      665 CGGATTTGAGTCAACAGATTTGATGAGCAATTAGAGAGATTAACCCAGCTTGTGTCTGG 724
Qy      818 AGCTACTTGGCTTACCGCTTGTGATGATTAACGCTGCAAAAGCACTAAATGTTTAAAGCG 877
Db      725 AACTTGTAGCTTGGCCCTTGTGATGATGATGCAATGCAAAAGAGCGAGGCTTTATG 784
Qy      878 GTGTGCGGAATATTTTGTGTGTGTGG 905
Db      785 GTGTACCGCAACATATATATGAGCTGTGG 812

```

```

RESULT 8
CD573714      897 bp      mRNA      linear      EST 12-JUN-2003
DEFINITION   UCRPT01_01_P12_T3 Poncirus trifoliata CTV-challenged cDNA library -
ACCESSION   CD573714
VERSION      CD573714.1 GI:31669616
KEYWORDS     EST.
SOURCE       Poncirus trifoliata
ORGANISM     Poncirus trifoliata
REFERENCE    1 (bases 1 to 897)
AUTHORS      Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D.,
              Mananaker,S., Choi,Y. and Kingan,T.
TITLE        Development of EST Resources and New Genetic Markers for California
              Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
              Unpublished (2003)
JOURNAL      Contact: Mikeal Roose
              Department of Botany & Plant Sciences, University of California
              Riverside, CA, 92521-0124, USA
              Tel: 9097874137
              Fax: 9097874437
              Email: mikeal.roose@ucr.edu
              Seq primer: T3.
FEATURES
  source
    1. 897
    /organism="Poncirus trifoliata"
    /mol_type="mRNA"
    /cultivar="Pomeroy OP"
    /db_xref="taxon:37690"
    /clone="UCRPT01_01_P12"
    /issue_type="Phloem"
    /dev_stage="10 - 30 cm shoots"
    /lab_host="E. coli TJC121"
    /clone_lib="Poncirus trifoliata CTV-challenged cDNA
    library - UCR"
    /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
    Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
    greenhouse at University of California, Riverside. The

```

ORIGIN

```

Query Match      17.8%; Score 429.4; DB 6; Length 897;
Best Local Similarity 70.4%; Pred. No. 7.5e-110;
Matches 574; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

scion was a open-pollinated (very probably selfed) seedling of poncirus trifoliata cv Pomeroy that was selected as homozygous for the Ctv resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the Ctv resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the Mt. Roose lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the University of California Riverside Genomics Institute, Core Instrumentation Facility. (Choi, Kingan).

Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Mananaker) using the HarvBST pipeline (http://harvest.ucr.edu) to remove contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank.

```

Qy      742 GCAAGTACCTTGCACCGGATTTTACGTCACAAATTTGATGAGACTTTGAGAGATCACT 801
Db      80 GCGAGTATGCTCCCTCCAGATTTTGACGACCAATTTGACAGAGTTCAGTCAAC 139
Qy      802 CCGCGTTATATGCTTGGAGTACTTGGCTTACCGCTTGTGATGATTAACGCTGCAAAAGA 861
Db      140 CCAAGCTGTCTTGTGAACTTTTAAAGCTTTTACCTTATGTGTATGATCAAGCAAGAGA 199
Qy      862 CTAATGTTTAAAGCGGTGTGCGAAATATTTTGTGTGTGTGAGAGAGGTGAGATCA 921
Db      200 GAAGAAGAGACTTATGATGATGATCAACATCTATGAGGCTTTGAGAGAGGAGGACAT 259
Qy      922 GCTCTTGTGTGGGCTTTGACCGCTGAGAACTTTATGATGAGCGCTTTTACGAATGACA 981
Db      260 GCAATTTGCTGGGGAATTTACACGTGAATCTTATGATGAGCGCTTTTACGATGACA 319
Qy      982 GCTGCTGAGAGGTTGATCTTTTGTAGTACCCCAAGCAATTTCCAGAGATCACTT 1041
Db      320 TCAAGCTGAGAGGTTTAAATATTTTGTGACACCAAAATGCAATCCAGCTGAACTTT 379
Qy      1042 GAAGTTTACGAATGCTACTTGTCTTGTGCTCAAGCTTTTATGTTAGAACCAAC 1101
Db      380 GAAGCTTATGAGATGAGCACTTGTACCTTGTGCCAAGCCTTTGTGGTAAAGCAACGCAT 439
Qy      1102 CTTTACAGAGATCTGATTAAGCAATTTCCAGAACTTCCAGAGCTTAAAGTATGCTATG 1161
Db      440 CTGATTTGACAGATCTGATTAAGCAATTTCCAGAACTTCCAGAGATTAAGTATGCAAGCTTT 499
Qy      1162 GAGATTTCCGAGATGTTTATGATGATACAGGAATTTGGAGATGACTTCCGCTTAA 1221
Db      500 CGGAGCTTGGCTCTATCTATATTTCTTGTGAAAACATGAGATGAGTTTGGATTTGAAA 559
Qy      1222 AGGGAGCTGTGCTACTGCTTATTAAGCAAGTTGATGAATCCGATATGTTGGCTTAA 1281
Db      560 AGGGGCTATATGTTCACTGCTTGAAGTAAAGCTTGAAGTGTGCTTATGTTGGCTTAA 619
Qy      1282 GACAGTGAAGATTCACAAATTTAGAGATCCAGATATGTTGAGTGTGTTTGGAGATTTCA 1341
Db      620 GATATGATTAAGTCCCATTAACAAATTTCCAGCATTTGATTTTGTGAAAACCTCA 679
Qy      1342 AATGATGATGACAAATGATGATCTCCCTGAGACTATGCAAAATGTTGAAAACCTGTTGGCA 1401

```

Db 680 AAGGAGCTAATGACATATGCTCTGAGCTGTAGAGATGGCTGGCA 739
 Qy 1402 GGGGTGCTCTTCTCTGCTTACAGACACCAAGATTTAAATTCGGGACTAC 1461
 Db 740 GAGGTGTTTCCCTGATTCAGGACACCAAGATTTAAATTCGGGACTAC 799
 Qy 1462 TATGATGATCTTATGTTTGAAGTTACTTGAAGAAGTGAAGTGAAGTTCCT 1521
 Db 800 TATGATGATCTTATGTTTGAAGTTACTTGAAGAAGTGAAGTGAAGTTCCT 859
 Qy 1522 TTAGCTGCTGCTCACTATGCAAGATTTGAGC 1556
 Db 860 TTAGCTGCAAGCTGCAATGATTTAGGCTTGAGC 894

RESULT 9
 B2437564/c 785 bp DNA linear GSS 13-DEC-2002
 LOCUS BONNR72TR_BO_1_6_2_KB.tot Brassica oleracea genomic clone BONNR72,
 DEFINITION genomic survey sequence.
 ACCESSION B2437564
 VERSION B2437564.1 GI:26691135
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 785)
 Ayele,M., Haas,B.J., Kumar,N., White,O.R., Van Aken,S.,
 Utterback,T.R., Mortman,J.R., White,C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its
 application to gene discovery and annotation in Arabidopsis
 Genome Res. 15 (4), 487-495 (2005)
 JOURNAL PUBMED 15805490
 COMMENT Other GSSs: BONNR72TF
 Contact: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..785
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BONNR72"
 /clone_1lb="BO_1.6_2_KB.tot"
 /note="Vector: pHD01, Site 1: BstXI, 1.6-2 kb sheared
 total DNA inserted into pHD01 using BstXI linkers"

ORIGIN
 Query Match 15.0%; Score 360.8; DB 9; Length 785;
 Best Local Similarity 86.4%; Pred. No. 2,1e-90;
 Matches 42; Conservative 0; Mismatches 57; Indels 10; Gaps 2;

Qy 505 AAGGTTCTGGGGCTCTCTGTATTTGCAAGAGTGTGAGACTGATAGTTCTTCGG 564
 Db 693 AAGGTTCTGGGGCTCTCTGTATTTGCAAGAGTGTGAGACTGATAGTTCTTCGT 634
 Qy 565 GTTGTAGAGCTGCTTAAAGAGAGGTTGCTTAAGCTTTAAGCAAGATGCTTTTA 624
 Db 633 GTAGGAAAGCTTGTCTTAAAGAGAGGTTGCTTAAAGCTTTAAGCAAGATGCTTTTG 574
 Qy 625 GTTATGCGCTTGGCTTCTGATGCTCGAGGAGTCTATGAGTTCACCTGAT 684
 Db 573 GTTATGCTGCTTGTCTTGTGATCTCCAGGATGCAATGCGATTAGATCTCTGAT 514

Qy 685 TTTATTAAGTGTATGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGCA 744
 Db 513 TTTATTAAGTGTATGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGCA 454
 Qy 745 AGTAGCTTGCACCGGATTTAGCTGACAAATTGATGAGACTTTGGAAGATTCCTCG 804
 Db 453 ACCAGCTTGCACCGGATTTAGCTGACAAATTGATGAGACTTTGGAAGATTCCTCG 394
 Qy 805 C-GTTATGCTGAGAGCTACTTGGCTTACCGCTTGTGATGATTCGCTCGAAAAAGCT 863
 Db 393 CGATTATGATTTGAGACTCTTGTGATTCCTTGGGGATGAT-----AAAGCA 343
 Qy 864 AATGTTTAAAGCGGTGCGGAATATTTGTCTGTTGAGAGAGTGAAGCATCAGC 923
 Db 342 AGATGTTTAAAGCGGTGCGGAATATTTGTCTGTTGAGAGAGTGAAGCATCAGC 283
 Qy 924 TCTTGTGGGGGTTTGAACCCGTGAGAAATTATGATGAGGCTTTTACGAATGACAC 983
 Db 282 TGTGTGTGGGTCTTACACAGTGAATTTATGATGAGGCTTTTACGAATGACAC 223
 Qy 984 TGCTGACAGGT 995
 Db 222 ATCTGAGCAAGT 211

RESULT 10
 CX195205 624 bp mRNA linear EST 28-DEC-2004
 LOCUS 13-E019365-021-009-J03-T7R ADIS-MP1Z 021 Brassica napus cDNA clone
 DEFINITION MP1Zp1022J039Q, mRNA sequence.
 ACCESSION CX195205
 VERSION CX195205.1 GI:56842629
 KEYWORDS EST.
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 624)
 Jakoby,M., Lehmann,D. and Weishaar,B.
 direct submission to GenBank (ADIS-MP1Z 021)
 Unpublished (2004)
 Contact: Bernd Weishaar
 Chair of Genome Research
 Bielefeld University, Institute for Genome Research
 Universitaetsstrasse 25, D-33594 Bielefeld, Germany
 Tel: +49-521-106 6873
 Fax: +49-521-106 6423
 Email: bernd.weishaar@uni-bielefeld.de
 Insert Length: 624 Std Error: 0.00
 Plate: 9 row: J column: 3
 Seq primer: T7R CTATATCGACTCATATAGGA.
 Location/Qualifiers
 1..624
 /organism="Brassica napus"
 /mol_type="mRNA"
 /culivar="Express 617"
 /db_xref="GABI:1111775"
 /db_xref="taxon:3708"
 /clone="MP1Zp1022J039Q"
 /tissue_type="flowers and flower buds"
 /dev_stage="young flowers and flower buds"
 /clone_1lb="ADIS-MP1Z 021"
 /note="Vector: pSPORT1, Site 1: SalI, Site 2: NotI; cDNA
 library from Brassica napus, strain Express 617; RNA was
 from young flowers and flower buds of two greenhouse
 plants; library was made at the Max-Planck-Institute for
 Plant Breeding Research, Cologne, Germany; cloning sites
 SalI-NotI, primer sites and orientation:
 T7-SalI-CCACGCTCCG-58rr-cDNA-polyA-CC-NotI-SP6; made by
 Marc Jakoby 09/2000; PI: Bernd Weishaar. Sequence
 submission managed by RZPD/GABI-Primary database:

ORIGIN

http://gabi.rzpd.de"

Query Match 14.9%; Score 359.2; DB 8; Length 624;

Best Local Similarity 76.9%; Pred. No. 5.6e-90;

Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;

1303 AGGAATCCAGCTATGTTGAGTTGTTTGGAGAAATCAATCGATGATGATGAT 1362

2 AGAATCCAGCTATGTTGAGTTGTTTGGAGAAATCAATCGATGATGATGATGAT 61

1363 CTCCTGAGCTATGCAAAATGTTGGAACCTGTTGCAAGGCTTGTCTTCTAGGTC 1422

62 CTCCTGAGCTATGCAAAATGTTGGAACCTGTTGCAAGGCTTGTCTTCTAGGTC 121

1423 AGAGACCCAAAGTAAATTTAAATCTGGGAGCTATGATGATGATGATGATGAT 1482

122 AGAGATACCAAGTAAATTTAAATCTGGGAGCTATGATGATGATGATGATGAT 181

1483 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1542

182 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 241

1543 GCAAGATGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1602

242 GCAAGATGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 301

1603 CTTTCCGCTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1662

302 CTTTCCGCTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 337

1663 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1719

338 AGTACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 397

1720 GCAGAGCTGTAAGACCTGTAAGACCTGTAAGACCTGTAAGACCTGTAAGACCT 1779

398 GCAGAGCTGTAAGACCTGTAAGACCTGTAAGACCTGTAAGACCTGTAAGACCT 431

1780 GTCTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1839

432 GTCTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 487

1840 GAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1899

488 GAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 547

1900 CAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1953

548 CAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 607

1954 TCTATGGAATCTGA 1967

608 TCTATGGAATCTGA 621

RESULT 11

LOCUS

CO075595

DEFINITION GR_Ea36G17.r GR_Ea Gossypium raimondii cDNA clone GR_Ea36G17.3',

mRNA sequence.

ACCESSION CO075595

VERSION CO075595.1

KEYWORDS GI:48745076

SOURCE EST.

ORGANISM Gossypium raimondii

Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Rosidae; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 769)

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,

Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and

Wing, R. A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: rwing@genome.arizona.edu

Plate: 36

row: G

column: 17.

FEATURES

source

1..769

Location/Qualifiers

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ea36G17"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH10B"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; Library made by Invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Clonies

plated/picked by Agt. More glycerol clones held in -80."

ORIGIN

Query Match 14.3%; Score 344.2; DB 7; Length 769;

Best Local Similarity 68.9%; Pred. No. 1.1e-85;

Matches 491; Conservative 0; Mismatches 213; Indels 9; Gaps 1;

1 ATGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60

54 ATGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 113

61 CCGGCGACGACAAAGCTCCGACGACGACGACGACGACGACGACGACGACGACGAC 120

114 CACCTCTCTTAAACCTTCACGACGACGACGACGACGACGACGACGACGACGACGAC 173

121 AATGGGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 171

174 AATGGGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 223

172 TCCT 231

234 TCCT 233

232 GATGCTCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 291

294 TCCT 353

292 ACACATTTCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351

354 ACTCATTTCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413

352 CAATTCGTTTGAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 411

414 CAATTCGTTTGAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 473

412 GAAACTCTGCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471

474 GAAACTCTGCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 533

472 GAAACTCTGCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531

534 CAGTACATCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593

532 CAGTACATCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591

594 CAGTACATCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653

592 TTGCTTAAGTCTTAAAGAGTGTGTTTAAAGTGTGTTTAAAGTGTGTTTAAAGTGTGTC 651

654 TTGCTTAAGTCTTAAAGAGTGTGTTTAAAGTGTGTTTAAAGTGTGTTTAAAGTGTGTC 713

Dept of Genetics and Biochemistry

Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA

Tel: 864 656 3060

Fax: 864 656 6879

Email: aalbert@clemson.edu

Total High Quality bases = 523

Seq primer: TAATACGACTACATACATAGG

High quality sequence stop: 631.

FEATURES

Source

1. 631

/organism="Prunus persica"

/mol_type="mRNA"

/culivar="Loring"

/db_xref="taxon:3750"

/clone="PP_LBA0027104f"

/tissue_type="Mesocarp"

/lab_host="E. coli"

/note="Vector: pBluescript II SK(-); Site 1: EcoRI;

Site 2: XhoI; authority: Prunus persica L. Batsch; The

sequence has been trimmed to remove vector sequence and

contains a minimum of 100 bases of phred value 20 or

above. For more details on library preparation and

sequence analysis go to

http://www.genome.clemson.edu/projects/peach.

To order

this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match

13.7%; Score 328.8; DB 5; Length 631;

Best Local Similarity

71.5%; Pred. No. 2,4e-81;

Matches

451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;

QY 913 GAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGTTTATGAATGAGCGCTTTTGA 972

DB 1 GCAGTTGCAATGTGGGGGNGATTCATACGTGAAATTTTCATGAAAGAGCGCTTCTTG 60

QY 973 CGAATGACAGCTGTGAGAGGTTGATCTTTTGTACTACCCCAAGCAATATTCACGA 1032

DB 61 CATATGCTGACGTGAGAGGTTGATTTATTTGTACTACCCCAAGCAATATTCACGA 120

QY 1033 GAGTCATTTGAGTTTACGAAGTTGCACTTCTTGTGCTCAAGCTTTTATTTGTAAG 1092

DB 121 GAAAGCTTTGAGTTTATGAGGTCCTTGTGCTTGTCTTCAAGCTTTTGTGTAAG 180

QY 1093 AAGCCACACCTTTTACGAGTGTGATGATGCAATTCAGCAACTTCAGCAGCTTAAGTA 1152

DB 181 AAACCTCATCACTTCAAGATGCTGAAACCTATTCAGAAACTTCAGCAGCTTAAGTA 240

QY 1153 ATGGCTATGAGATTCCTGAGATTTGTATGATACAGGAATTAATGGGATGATGATTC 1212

DB 241 ACAGCTGTAGACATTTCTTGAACACTATATTAACAAAGAGCACTGATGATGATTC 300

QY 1213 GGTCTGAAAGGGGACTCTGTGACATGCTTATAGGCAAGTTGATGATACCGTATGAG 1272

DB 301 GCTTTGAGAGGGGACTCTGTGACATGCTTATAGGCAAGTTGATGATACCGTATGAG 360

QY 1273 TTGGGCTTACAGTGAAGATTCACATATGAGAAATTCAGCTATTTGAGTTTGTGG 1332

DB 361 TTGGGCTTACAGTGAAGATTCACATATGAGAAATTCAGCTATTTGAGTTTGTGG 420

QY 1333 GAAATTTCAATGATGATG-----ACATGATGATGCTCCCTGAGCTATGAGAA 1380

DB 421 GAAATTTCAATGATGATG-----ACATGATGATGCTCCCTGAGCTATGAGAA 480

QY 1381 TTGTTGAAACCTGTTGAGAGGGTGTCTTTCTTGTGTTGAGAGACCAAGATTA 1440

DB 481 CTATTTGAGAGCTGTTGAGAGGGTGTCTTTCTTGTGTTGAGAGACCAAGATTA 540

QY 1441 AAATTTAACTCGGGGACTATGATGATCTATGATGTTTGTGTTGTTGTTGTTGTTG 1500

DB 541 GAGTTTCAGACTGGAGACTATGATGATCTATGATGATCTTGTGATGATCTTGAAGAGCTG 600

QY 1501 GAGTAGTTGAGGTTCTCTTACTGCTG 1531

DB 601 GATGCACTAATGTTCACTTACTGCTG 631

RESULT 14

CN918842

LOCUS

DEFINITION

ABPB006290, mRNA sequence.

ACCESSION

CN918842

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Source

1. 703

/organism="Malus x domestica"

/mol_type="mRNA"

/db_xref="taxon:3750"

/clone="ABPB006290"

/tissue_type="Root tips (distal 1.5 cm)"

/note="Vector: pBluescript SK(-); library sequenced by

Genes Research & Development"

ORIGIN

Query Match

13.6%; Score 326.6; DB 7; Length 703;

Best Local Similarity

71.5%; Pred. No. 1e-80;

Matches

450; Conservative 0; Mismatches 164; Indels 15; Gaps 1;

QY 74 AGCTCGACGTAGCCCAACCACTCTACATATCTGCTCCGACGAATGGCCGACC 133

DB 75 ACCACGGAAACCCCAAAAGCTCCCTCCCACTCTGCTCCGACGAATGGCCGACC 134

QY 134 GTCTTCTCTGCACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCT 181

DB 135 GCTCTCTCTGCACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCT 194

QY 182 ---TCGCAACGGCACACACCGCACTGCTGCTCTGCAACCACTATGATGCTG 238

DB 195 CCTCTCTCTGCACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCT 254

QY 239 CCGAAGCCCACTGCTCCCATCCCATGATTTCTACCAAGATTAAGAGCTCAACACAT 298

DB 255 CCGAAGCCCACTGCTCCCATCCCATGATTTCTACCAAGATTAAGAGCTCAACACAT 314

QY 299 TCTTAACCGATGAAATCAGAAAGCATTTGAAAGCTAGGGTTTGAACCGCCCAATTG 358

DB 315 TCTTAACCGATGAAATCAGAAAGCATTTGAAAGCTAGGGTTTGAACCGCCCAATTG 374

QY 359 GTTTAGCGACGACGCTTTAATCAGCGGAGACAGATTTCTTAAGCTGTTGGAACCTG 418

DB 375 GTTTAGCGACGACGCTTTAATCAGCGGAGACAGATTTCTTAAGCTGTTGGAACCTG 434

QY 419 TGTCTATCTCGGTCTAGAAAGAGTACATGAGGCTTTCTTGTATGATGAAGAGCTA 478

DB 435 TAGCCGACCCCGCTCCAGAAAGAGTACATGAGGCTTTCTTGTATGATGAAGAGCTA 494

QY 479 CAGTCATCACTGATGTTCTTGGGATPAAGGTTCTGCGGCTCTCTGNTATTGCAAG 538
 DB 495 CCATCATCACTCAAGTTCTTGGGATPAAGGTTCTGAGCTCTGTGCTCTGCAAG 554
 QY 539 GTGGTGAAGTGAATGTTCTTGGGATPAAGGTTCTGAGCTCTGNTATTGCAAG 558
 DB 555 CTGGGAAGTGAAGTGAATGTTCTTGGGATPAAGGTTCTGAGCTCTGNTATTGCAAG 614
 QY 599 AGTCGTTTAAAGCAAGATGTTTAAAGTGAAGGTTCTGAGCTCTGNTATTGCAAG 658
 DB 615 AGTCGTTTAAAGCAAGATGTTTAAAGTGAAGGTTCTGAGCTCTGNTATTGCAAG 674
 QY 659 ATGCTATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
 DB 675 ATGCTATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 703

RESULT 15
 BP785511/c 415 bp mRNA linear EST 10-FEB-2005
 LOCUS BP785511 RAFL7 Arabidopsis thaliana cDNA clone RAFL07-95-D11 3',
 DEFINITION mRNA sequence.
 ACCESSION BP785511 GI:59261187
 VERSION BP785511
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 415)
 REFERENCE
 AUTHORS Seki, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K.,
 Iida, K., Enju, A., Sakurai, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and
 Shinzaki, K.
 Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
 Unpublished (2005)
 COMMENT
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998, 2002). This clone is in a
 modified pBluescript vector.
 Please visit our web site (http://pfweb.gsc.riken.jp and
 http://large.gsc.riken.jp) for further details.
 reversed clone.

FEATURES
 source
 Location/Qualifiers
 1..415
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL07-95-D11"
 /dev_stage="rosette plants"
 /lab_host="DH10B"
 /clone_lib="RAFL7"
 /note="Site 1: BamHI; Site 2: SalI; subjected to
 cold-created (1, 2, 5, 10, 24 hr)"

ORIGIN
 Query Match 13.3%; Score 321; DB 3; Length 415;
 Best Local Similarity 100.0%; Pred. No. 3.4e-79;
 Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2086 CCTGATTCACCGCATAGAAATGTTTACAGAGTTTGGATGGCGCAATGCTGAAGATTGG 2145
 DB 415 CCTGATTCACCGCATAGAAATGTTTACAGAGTTTGGATGGCGCAATGCTGAAGATTGG 356
 QY 2146 ACTGACAGAGCAGCTGAATGCGCAGCTGGTGGTTATGATATACACTGTGAAA 2205

DB 355 ACTGACAGAGCAGCTGAATGCGCAGCTGGTGGTTATGATATACACTGTGAAA 296
 QY 2206 CTATCTGTGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2265
 DB 295 CTATCTGTGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 236
 QY 2266 CTGGAAGAGTCTGCTTGTCTATCTGAATTTGGTTTCATCCAGAAAACAATGCTACTGATGTC 2325
 DB 235 CTGGAAGAGTCTGCTTGTCTATCTGAATTTGGTTTCATCCAGAAAACAATGCTACTGATGTC 176
 QY 2326 AGAAGCTTACCAAGATTCAGAAATTTCTGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 2385
 DB 175 AGAAGCTTACCAAGATTCAGAAATTTCTGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 116
 QY 2386 GGCTCTGTTCTGTCATCATTA 2406
 DB 115 GGCTCTGTTCTGTCATCATTA 95

Search completed: February 21, 2006, 15:40:23
 Job time : 6314 secs